1993-03

Ensemble forecasting techniques in medium-range forecasting

Warren, Steven W.

Monterey, California. Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

ENSEMBLE FORECASTING TECHNIQUES
IN
MEDIUM-RANGE FORECASTING

by

Steven W. Warren

March, 1993

Thesis Advisor: Wendell A. Nuss

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL (*If applicable*) 35	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (*City, State, and ZIP Code*) Monterey, CA 93943-5000	7b. ADDRESS (*City, State, and ZIP Code*) Monterey, CA 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*)	10. SOURCE OF FUNDING NUMBERS Program Element No. / Project No / Task No. / Work Unit Accession Number
11. TITLE (*Include Security Classification*) ENSEMBLE FORECASTING TECHNIQUES IN MEDIUM-RANGE FORECASTING	
12. PERSONAL AUTHOR(S) Warren, Steven W.	
13a. TYPE OF REPORT Master's Thesis 13b. TIME COVERED From To	14. DATE OF REPORT (*year, month, day*) 1993, March 15. PAGE COUNT 126
16. SUPPLEMENTARY NOTATION The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.	
17. COSATI CODES FIELD / GROUP / SUBGROUP	18. SUBJECT TERMS (*continue on reverse if necessary and identify by block number*) Ensemble models, regression technique, forecast divergence, systematic error

19. ABSTRACT (*continue on reverse if necessary and identify by block number*)

A continuing trend in numerical weather prediction (NWP) is the desire for reduced model forecast error. Developments in NWP such as advanced computing power and improved model physics and analysis methods have been successful in lowering error but are potentially limited. The regression method of ensemble forecasting is used to further reduce mean forecast error when compared to individual model forecast performances. A statistical regression scheme is utilized to achieve an optimum combination fitting of the National Meteorological Center, the European Centre for Medium-Range Weather Forecasts, and the U.S. Navy Fleet Numerical Oceanography Center forecast models. The performance of the regression model is evaluated for 72-h and 108-h prediction cycles through statistical and subjective comparisons with the individual models and an equally weighted ensemble model at the surface and at 500 hPa. The regression model is shown to produce significant gains through the reduction of systematic error present in the individual model forecasts.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS REPORT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Wendell A. Nuss	22b. TELEPHONE (*Include Area code*) (408) 656-2308 22c. OFFICE SYMBOL MR/Nu

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Approved for public release; distribution is unlimited.

Ensemble Forecasting Techniques
in Medium-Range Forecasting

by

Steven W. Warren
Lieutenant, United States Navy
B.S., University of Oklahoma, 1983

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY AND PHYSICAL OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
March 1993

Author: [redacted]
Steven W. Warren

Approved by: [redacted]
Wendell A. Nuss, Thesis Advisor

[redacted]
Carlyle H. Wash, Second Reader

[redacted]
Robert L. Haney, Chairman
Department of Meteorology

ABSTRACT

A continuing trend in numerical weather prediction (NWP) is the desire for reduced model forecast error. Developments in NWP such as advanced computing power and improved model physics and analysis methods have been successful in lowering error but are potentially limited. The regression method of ensemble forecasting is used to further reduce mean forecast error when compared to individual model forecast performances. A statistical regression scheme is utilized to achieve an optimum combination fitting of the National Meteorological Center, the European Centre for Medium-Range Weather Forecasts, and the U.S. Navy Fleet Numerical Oceanography Center forecast models. The performance of the regression model is evaluated for 72-h and 108-h prediction cycles through statistical and subjective comparisons with the individual models and an equally weighted ensemble model at the surface and at 500 hPa. The regression model is shown to produce significant gains through the reduction of systematic error present in the individual model forecasts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. THE SIGNIFICANCE OF ENSEMBLE FORECASTING

A continuing trend in numerical weather prediction (NWP) is the desire for reduced model forecast error. A number of developments in NWP have produced notable success in reducing forecast error in numerical models. These developments include improved model physics, dynamics, increased model resolution, improved analysis and data assimilation methods, and advances in computing power.

Although substantial success has been achieved, model predictability limits remain and are created by the growth in time of small errors in the initial state of the atmosphere utilized in model predictions. Additionally within individual models, systematic errors can reduce the accuracy of model forecasts. Statistical approaches have been presented which address these errors to a degree. One notable approach was the development of model output statistics (MOS) (Glahn and Lowry, 1972).

Today with the development and availability of various numerical model products, any given forecast office may obtain several different sets of NWP forecasts as prediction tools. Although originating from different models, forecasts of like parameters on coincident prediction cycles are often

available. These various model forecasts of the same parameter may, however, produce significantly different results at times. Therefore, the individual forecaster is presented with the problem of choosing between individual forecasts or attempting to consolidate or combine the forecasts in some manner to achieve an improved ensemble forecast.

Operational forecasters have qualitatively combined different model forecasts for years based on incorporating biases determined from published model performance statistics and on personal experience in working with the models. Quantitative combinations have also been developed by forecasters and have been gaining in sophistication with the advent of increased computing power. These more formal quantitative or statistical treatments of various forecast models have been shown to produce successful results in improving forecast accuracy in ensemble forecast methods when compared to individual forecasts.

The ability to forecast the skill of these improved forecasts has also been an area of significant research in recent years. Prediction of forecast skill is seen by many to be as important as the forecast itself, given the variance in geographical region, seasonality, and atmospheric and boundary states from one forecast case to another. Consequently, successfully linking a reliable forecast skill prediction capability with a tested ensemble forecast model would provide

the operational forecaster with an added tool to reduce overall model forecast error and increase user confidence in the skill level of a prepared forecast.

B. OBJECTIVE

The goal of this study is to investigate the utility of combining several different numerical prediction models through a multiple linear regression statistical method to produce an improved forecast over a medium range prediction cycle. The study first utilizes 72-h forecasts and later also includes an investigation of the use of longer range forecasts. The forecast models include the National Meteorological Center (NMC) aviation run (AVN) and medium-range forecast (MRF) models, the European Centre for Medium Range Weather Forecasts (ECMWF) model, and the U.S. Navy Fleet Numerical Oceanography Center (FNOC) forecast model.

A statistical regression scheme will be utilized to achieve an optimum combination fitting of the forecast models for both 72-h and extended range forecasts. The desired result will be a combined forecast showing reduced model forecast error over individual model forecasts.

Section II of this thesis provides background on previous studies conducted in the areas of ensemble forecasting and the prediction of forecast skill. Section III discusses data analysis. Section IV provides several case studies investigating the relationship between individual model

forecasts and resultant ensemble forecasts for a region encompassing the eastern Pacific Ocean and the western United States. The final section contains conclusions and recommendations.

II. BACKGROUND

A. PREVIOUS STUDIES

The central goal of many of the previous studies in the area of improved NWP skill has been to maximize forecast skill and to demonstrate the ability to predict skill a priori. The ensemble forecasting technique developed by Leith (1974) and Hoffman and Kalnay (1983) was a basis for much of this work. Leith developed the Monte Carlo forecasting (MCF) technique which involves running a number of model integrations from distinct initial states. The initial states can be generated by adding a series of perturbations to the latest operational analysis. Hoffman and Kalnay developed the lagged average forecasting (LAF) method which used initial states generated by using a series of operational analyses, each lagging the most recent by a different time increment. As in the MCF technique, a number of model integrations were run from these distinct initial states. The perturbations generated by the LAF method have an advantage over the MCF technique since they are based on governing dynamics (Dalcher et al., 1988).

A number of studies have branched from these previous works of Leith and Hoffman and Kalnay. Kalnay and Dalcher (1987) utilized the LAF method but only as a background in studying the predictability of forecast skill for a forecast

model on a regional basis. This was a refinement to a similar study on skill predictability by Dalcher et al. (1985) which made use of global vice regional verifications. The regional study demonstrated that the divergence of the forecast values of ensemble members from ensembles as small as three or four members can be a good a priori prediction of forecast skill. Dalcher et al. (1988) achieved encouraging results with the LAF method in medium-range forecasts by using a weighting scheme based on horizontal wavenumbers. Recently, Murphy (1990) used LAF experiments to investigate benefits arising from the use of ensemble forecasts in extended range forecasts out to a month time frame.

The trend toward follow-on studies focusing on the idea of forecasting forecast skill includes such research as Leslie et al. (1989). This work demonstrated that the skill of short-term regional numerical forecasts can be predicted on a day-to-day basis. Their predictions were achieved by using a statistical regression scheme with the model forecast errors as predictands and the initial analysis, together with the model forecast, at proximate points, as the predictors. Leslie and Holland (1991) also studied the predictability of regional forecast skill using both single and ensemble forecast techniques. These techniques included correlating model forecast error with the divergence or spread of ensemble member forecasts, and with a predictor based on a statistical regression scheme developed by Bennett and Leslie (1981) and

improved by Glowacki (1988). A final technique used by Leslie and Holland was the correlation of model forecast error with a persistence predictor method developed by Chen (1989). This method used the persistence of the model forecast vice the spread of ensemble members to predict forecast skill. The persistence of the model forecast within the latest integration was used as the predictor. This predictor showed success on both regional and hemispheric scales.

B. REGRESSION TECHNIQUE FOR ENSEMBLE FORECASTING

As mentioned, an operational forecaster may have several forecast models available to use in weather prediction. The forecaster ultimately desires to be able to combine these forecast tools to provide a single improved forecast product. The technique of multiple linear regression is one method that can be used to combine available forecast models. In this study, the NMC AVN and MRF, ECMWF, and FNOC models were chosen as the products to be combined through a multiple linear regression technique.

1. IMSL Statistics/Library (IMSL)

In order to develop a regression model from the above forecast products, subroutines were used from the IMSL, a commercial software library. These subroutines utilized grid data from the forecast models to formulate a multivariate regression model. A corrected sum of squares and crossproducts matrix was developed as an input for the

regression model and an intercept was included. The subroutines utilized the Cholesky factorization in regression computations.

The output from the IMSL subroutines consisted of the regression coefficient and intercept values which allowed for a fitted regression equation of the form $y=a_0+a_1x_1+a_2x_2+a_3x_3$. The values x_1, x_2, and x_3 represent forecast values from the ECMWF, NMC, and FNOC forecast models. The values a_1, a_2, and a_3 represent the regression coefficient values while a_0 represents the model intercept value. The y value is the final ensemble forecast value resulting from the fitted model equation.

2. Application for 72-h Forecasts

In applying the multiple regression techniques to the 72-h forecast model cycle, the 72-hr NMC aviation run (AVN) was used along with the 72-h forecasts from the ECMWF and FNOC models. The three forecasts were used as independent values while the ECMWF analyses valid at the 72-h verification times were used as the dependent values in the regression model. The reasoning for the selection of the ECMWF analyses as verification products is discussed briefly in Section 4 and more completely in Chapter 3.

3. Application for Extended Forecasts

The regression model was also applied to extended forecasts in an effort to determine whether the method could

be used to achieve gains over a longer forecast cycle. Application of the regression model to extended range forecasts required the use of the 120-h ECMWF forecasts along with the 108-h NMC MRF and FNOC forecasts to form what was essentially a 108-h ensemble forecast. One of the efforts of this study was to make use of the forecast products available at a given forecast time. As such, the use of a staggered set of extended forecasts was required. The MRF and FNOC products were available only on a 0000 UTC forecast run, while the ECMWF products were available only on the 1200 UTC forecast cycle. The staggered forecasts were therefore required to allow for a set of forecasts to be used in the regression model with a single coincident verification time. The three extended forecasts were used as the independent values in the regression model, while the ECMWF analyses valid at the extended range verification time were used as the dependent values.

4. Choice for Verifying Analysis

In order to determine which forecast center's analyses would be used as verification products in the regression scheme, an error analysis study was conducted using the NMC, FNOC, and ECMWF analyses and 72-h forecast products. The results will be discussed later in the data analysis section, however, the ECMWF analyses showed generally encouraging

results against each center's forecast products and was chosen as th verification product.

The ECMWF analyses and forecast products are highly respected by the NWP community. The data used in the ECMWF analyses include global satellite data (SATOB, TOVS, SATEM), global free-atmosphere data (AIREP, AMDAR, TEMP, PILOT), oceanic data (SYNOP/SHIP, PILOT/SHIP, TEMP/SHIP, DRIBU), and land data (SYNOP). Approximately 40,000 observational data points are used in each analysis and data checking and validation is applied to each parameter. The data utilized in the ECMWF analyses are similar to the data used in analyses from other forecast centers. The accuracy of these analyses, however, depends on the accuracy of a model's first guess. In this respect, the higher resolution of the ECMWF model may contribute to a higher accuracy in the ECMWF analyses.

III. DATA ANALYSIS

A. ANALYSIS METHOD

Surface and 500-hPa analyses and forecasts over a seven month period from April through December, 1992 were collected from products received at the Naval Postgraduate School's Meteorology Department for the study. The northern hemisphere analyses were utilized along with 72-h and 108-h FNOC Navy Operational Global Analysis and Prediction System (NOGAPS) and NMC Aviation (AVN) and Medium-Range Forecasts (MRF) as well as 72-h and 120-h ECMWF Global Spectral Model forecasts. The products from these centers were utilized due to their availability through similar forecast periods, their overlapping forecast and analysis model domains, as well as their wide usage and acceptance as medium-range numerical forecast guidance.

An error analysis was initially conducted to determine the relative performance of the model forecasts against both their own analyses and the analyses based on the other models. A regression analysis was then applied to a subset of the data, April through October, to obtain an optimum regression fitting of this three model combination. Upon review of the performance of the regression method within the dependent data

set, the method was then carried forward and applied to independent data from the November and December periods.

1. Model Forecast and Analysis Grid Modifications

In order to utilize the various model forecasts and analyses in the regression method, differing model grid resolutions as received at the Naval Postgraduate School had to be addressed. The ECMWF forecast and analysis grid resolution was 2.5° in both latitudinal and longitudinal directions. This compares with the NMC AVN which is received with a latitudinal grid resolution of 2.5° and a longitudinal resolution of 5.0°. The NMC MRF and FNOC products are both received with latitudinal and longitudinal resolutions of 5.0°. Since the FNOC products were common to both the 72-h and 108-h forecast studies, the remaining product grids were reduced in resolution to correspond with the overall 5.0 degree grid resolution. This grid resolution consisted of simply dropping grid points that did not match the lowest resolution grid. This grid resolution allowed the data from each model to be readily used in the regression calculations and error analysis studies.

2. Selection of Model Verification Analysis

A model verification study was conducted to determine the validity or "fairness" of choosing a single model analysis as the verification analysis for use in the regression equations over the remaining available analyses. The error

analysis was conducted using a straight differencing method of subtracting verifying analysis values from forecast values at individual grid points over a chosen domain. Each model's 72-h forecast for both the surface pressure and 500-hPa height was verified with the differencing method against the verifying analysis produced by the data assimilation using the same model and against the verifying analysis using the ECMWF model. The model forecasts were verified in this manner over a seven month period of data from April through October.

Average error values over the entire seven month period along with mean field values were plotted for each forecast and analysis combination over a northern hemispheric domain. Examples using the ECMWF forecasts and analyses are shown in Fig. 1 and Fig. 2, which show the mean seven month period errors of the surface (Fig. 1) and 500-hPa (Fig. 2) ECMWF forecasts overlaid upon corresponding mean surface pressure and 500-hPa height values.

As a first indication of model performance using a single analysis, the mean analyzed structure was compared to climatology and the mean model error based on the ECMWF analysis. At the surface, the main synoptic features indicated in the mean field values of surface pressure from April through October (Fig. 1) related fairly well in position and intensity to climatology. Expected climatological features included high pressure areas over the northeast Pacific Ocean and northern Atlantic Ocean, a monsoonal low

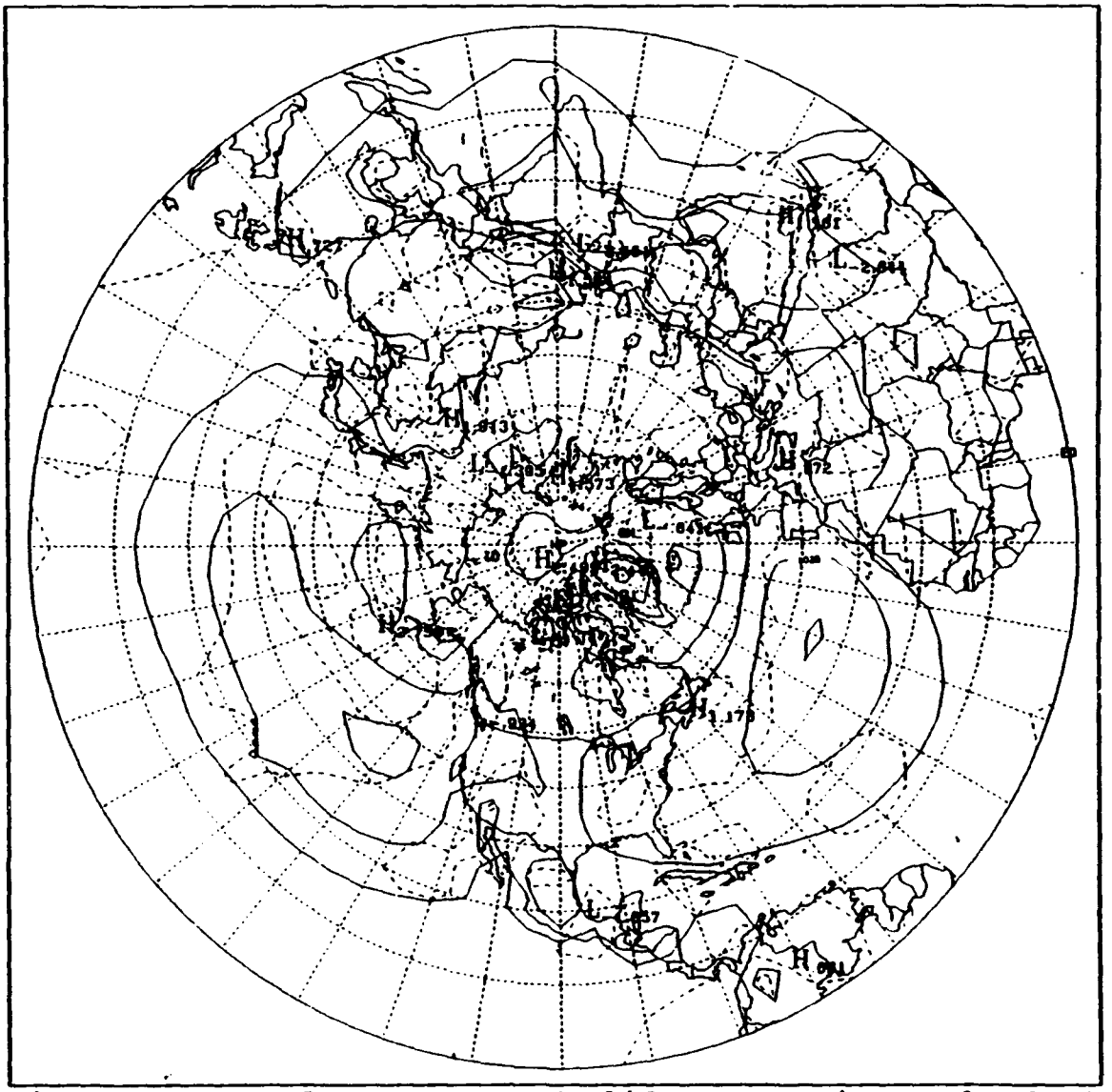

Fig. 1. Mean surface pressure (solid, contour interval 4 hPa) from ECMWF analyses and mean forecast error (dashed, contour interval 1 hPa) from ECMWF forecasts and analyses during April-October 1992.

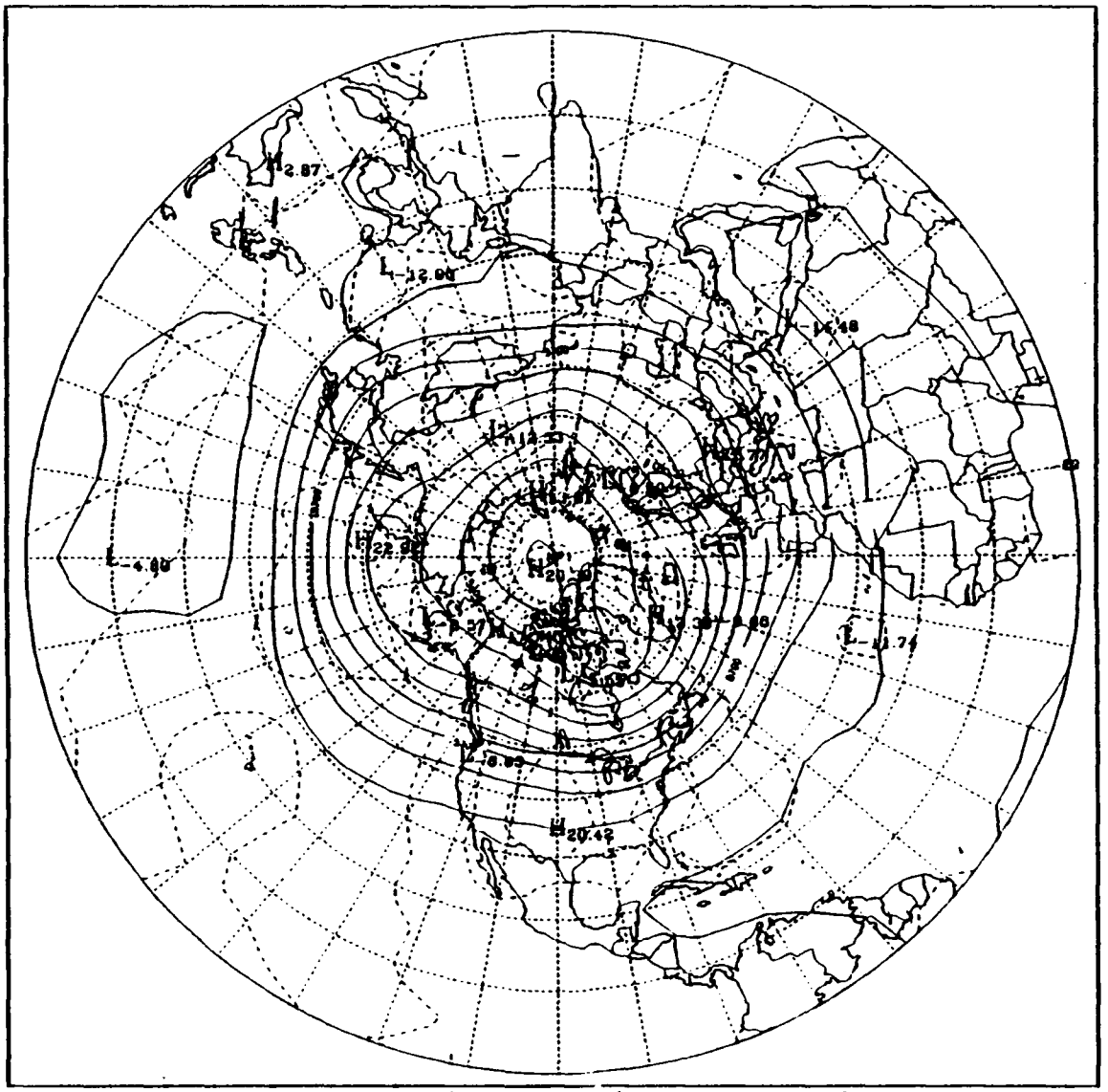

Fig. 2. Mean 500-hPa height (solid, contour interval 60 m) from ECMWF analyses and mean forecast error (dashed, contour interval 10 m) from ECMWF forecasts and analyses during April-October 1992.

pressure area extending from the Middle East across Asia, and a thermal low pressure area in the southwestern United States.

A review of the forecast performance by the individual model forecasts during the seven month period near these synoptic features indicated characteristic errors that occurred in each model's performance. For example, the mean forecast error of the ECMWF surface pressure forecasts (Fig. 1) was high in the north to northeastern areas of the subtropical high pressure regions in the northern Pacific and Atlantic Oceans. Additionally, relatively high errors were present in the ECMWF surface forecasts near the Himalayas and in the surrounding vicinity of highly variable terrain. The mean forecast errors of the NMC AVN and FNOC surface pressure forecasts based on the ECMWF analysis (not shown) exhibited high error patterns similar to the ECMWF forecasts in the oceanic subtropical high regions. In the Atlantic high pressure region, the NMC AVN and FNOC error patterns matched the ECMWF error pattern closely in position and were within approximately 0.3 hPa of the ECMWF error in magnitude. In the Pacific high pressure region, the NMC AVN and FNOC error patterns were again very similar in position to the ECMWF error. Differences in the magnitude of the error were present, however, with the NMC AVN and FNOC forecasts exhibiting error values approximately 0.4 hPa and 0.8 hPa higher, respectively, than the ECMWF error. The error patterns of the NMC AVN and FNOC surface pressure forecasts

near the Himalayas differed from the ECMWF forecasts both in orientation and magnitude, indicating that each of the products handled the terrain in this area somewhat differently. The NMC AVN forecasts exhibited the highest above and below zero errors in this region, with mean forecast errors of up to +5.226 hPa and -4.598 hPa relative to zero in the area.

At 500 hPa, the main synoptic features indicated in the mean height pattern from April through October were generally similar in position and magnitude to expected climatological features for this time period. Comparison of the mean forecast errors from the ECMWF (Fig. 2), NMC AVN (not shown), and FNOC (not shown) 500-hPa forecasts indicated that these products exhibited similar trends in forecasting the position of the main synoptic features at 500 hPa during the period but differed at times in forecast magnitude. For example, all three forecast products exhibited a relative maximum or minimum in mean forecast error in the lower portion of the trough extending northeast to southwest across the eastern United States. This is indicated by a high error value of +20.42 m on the ECMWF mean forecast error plot (Fig. 2). The NMC AVN and FNOC mean forecast error plots indicated extremes in error values at the same location, however, the error values for these products were -0.82 m and +18.81 m, respectively. Similarities in the errors of the three forecast models were also noted in the ridge located along the

west coast of the United States. All three products forecast height values too low in this region with varying degrees of magnitude. The mean forecast errors in this region for the ECMWF, NMC AVN, and FNOC forecasts were -6.95 m, -39.51 m, and -24.93 m, respectively.

Comparison of the mean forecast errors both at the surface and at 500 hPa indicated that the three forecast models generally exhibited similar characteristics in the distribution of forecast error when forecasting the positions of synoptic-scale features. Error characteristics did, however, differ at times in terms of magnitude for these features. However as noted in the region of the Himalayas, the model error characteristics tended to vary both in position and magnitude in regions of highly variable terrain. This indicates a variability in the treatment of terrain by the three models.

Along with the mean forecast error plots, overall mean error statistics over the grid were also calculated to present a single average error over the hemispheric domain and regional domains of 0°-90°N, 35°-90°N, and 50°-90°N for each forecast and analysis combination. The results are tabulated in Table 1 and Table 2. The error values were computed at each grid point and then averaged for each specific regional domain shown in the tables to obtain overall errors.

The comparison of surface FNOC forecasts with FNOC and ECMWF analyses in Table 1 showed that the FNOC products

TABLE 1. ERROR ANALYSIS OF 72-h SURFACE PRESSURE FORECAST VERIFICATION DURING APRIL-OCTOBER 1992.

FORECAST TYPE/ VERIFYING ANALYSIS	AVERAGE ERROR (mb) (FORECAST-ANALYSIS=ERROR) 0°-90°N	35°-90°N	50°-90°N
FNOC/FNOC	0.196	0.391	0.155
FNOC/ECMWF	-0.106	0.118	-0.006
NMC AVN/AVN	-0.136	-0.006	-0.170
NMC AVN/ECMWF	-0.373	-0.460	-0.588
ECMWF/ECMWF	0.107	0.439	0.536

performed better on average when verified against the ECMWF analyses than against the FNOC analyses. At 500 hPa, the FNOC forecasts performed slightly better, approximately three meters lower overall error, against the FNOC analyses than against the ECMWF products (Table 2). The error analysis for this data set shows that use of the ECMWF analysis as the verification product does not result in any serious disadvantages in terms of error statistics calculated for the FNOC forecasts. Since these differences were small, the use of the ECMWF analysis for verification was considered to be reasonably fair in this case.

The error comparisons for the AVN forecasts in Table 1 and Table 2 show that at the surface and 500 hPa the AVN model performs better against its own analyses than against the ECMWF products. The error verification against the AVN

TABLE 2. ERROR ANALYSIS OF 72-h 500-hPa HEIGHT FORECAST VERIFICATION DURING APRIL-OCTOBER 1992.

FORECAST TYPE/ VERIFYING ANALYSIS	AVERAGE ERROR (meters) (FORECAST-ANALYSIS=ERROR) 0°-90°N	35°-90°N	50°-90°N
FNOC/FNOC	-6.051	-8.498	-10.581
FNOC/ECMWF	-8.517	-11.565	-13.332
NMC AVN/AVN	-9 639	-4.852	-3.935
NMC AVN/ECMWF	-22 15	-17.459	-16.178
ECMWF/ECMWF	2.413	5.174	5.845

analysis was, on average, approximately 13 meters better at 500 hPa and up to approximately 0.4 hPa at the surface.

The error reduction in verifying the AVN forecasts against AVN analyses vice ECMWF analyses is more significant than that found for the FNOC products, which suggests that it is not fair to verify these products against the ECMWF products. However, the distribution of the mean errors for the AVN model verified against its own analysis and the ECMWF analysis, which was discussed above, were similar. This suggests that these differences represent systematic differences between the models, for which the ensemble regression forecast is expected to account.

3. Application of the Regression Method

Once the verifying analysis choice was made and the model grid resolution differences were resolved, the grid data was then available for use in the multiple linear regression

calculations. The regression was done independently at each grid point for the analysis period. While each point is not strictly independent, this method allows for regional variations in the performance of any particular model.

a. Regression Coefficients and Intercept Values

Daily NMC, FNOC, and ECMWF forecast values at the surface and 500 hPa over a seven month period from April through October were used as the independent values in the regression formulas. As mentioned, the ECMWF analyses at the surface and 500 hPa were chosen as the verification analyses and represented the dependent values in the regression formulas.

Regression fitting of the grid data was conducted for both the 72-h and extended range forecasts, yielding two separate sets of grid point regression equations based on forecast length. For the extended range forecast regression problem, NMC MRF and FNOC 108-h forecasts were used along with 120-h ECMWF forecasts as the independent values in the regression calculations. Although the forecast period differed between these model forecasts, the verification time was coincident for all of the forecasts. The use of different forecast periods was necessitated by the limited availability of the products at a given forecast time. The NMC MRF forecast products were available only for a daily 0000 UTC run while the ECMWF products were available only for a 1200 UTC

run. The FNOC products were also available for the 0000 UTC run and consequently, the use of these products from the 0000 UTC run required that the 120-h ECMWF forecast from the prior 1200 UTC run be used to allow for a single coincident valid time for forecast verification.

As previously mentioned, IMSL subroutines were used to calculate regression coefficients and intercept values at each grid point. Plots of the regression coefficient and intercept values for the 72-h and 108-h forecasts were then produced for the northern hemisphere. Plots of coefficient values for the surface forecasts are provided in Fig. 3, Fig. 4, and Fig. 5. Although a detailed investigation of these plots was not conducted, examples of preferential model weighting were identified in the individual model coefficient plots. For example, a higher weighting was evident in the AVN coefficients plot (Fig. 4) off the east coasts of Asia and the United States. Additionally, the ECMWF forecasts received a higher weighting in the central oceanic areas of the Pacific and Atlantic Oceans (Fig. 3). These examples suggested that the regression method was identifying model performance trends in specific regions. Identification of these trends is essential in enhancing the performance of the ensemble forecast blend by the regression method. Preferential weighting was again apparent in the 500-hPa coefficients plots (not shown), however, distinct trends in the weighting were not as obvious. The surface forecast intercept values (Fig.

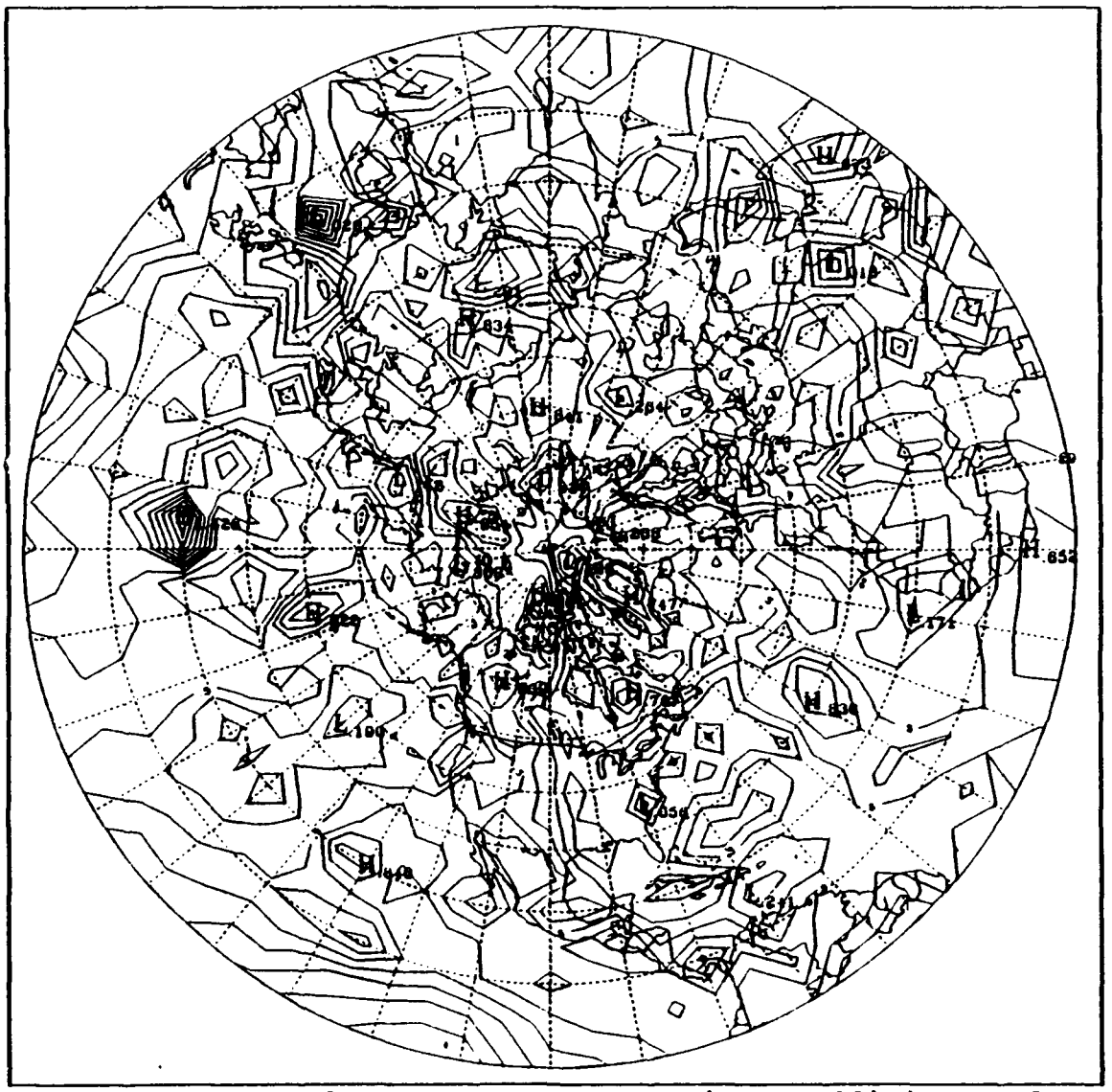

Fig. 3. ECMWF surface pressure regression coefficient values (solid, contour interval 0.1) for the dependent data set (April-October 1992).

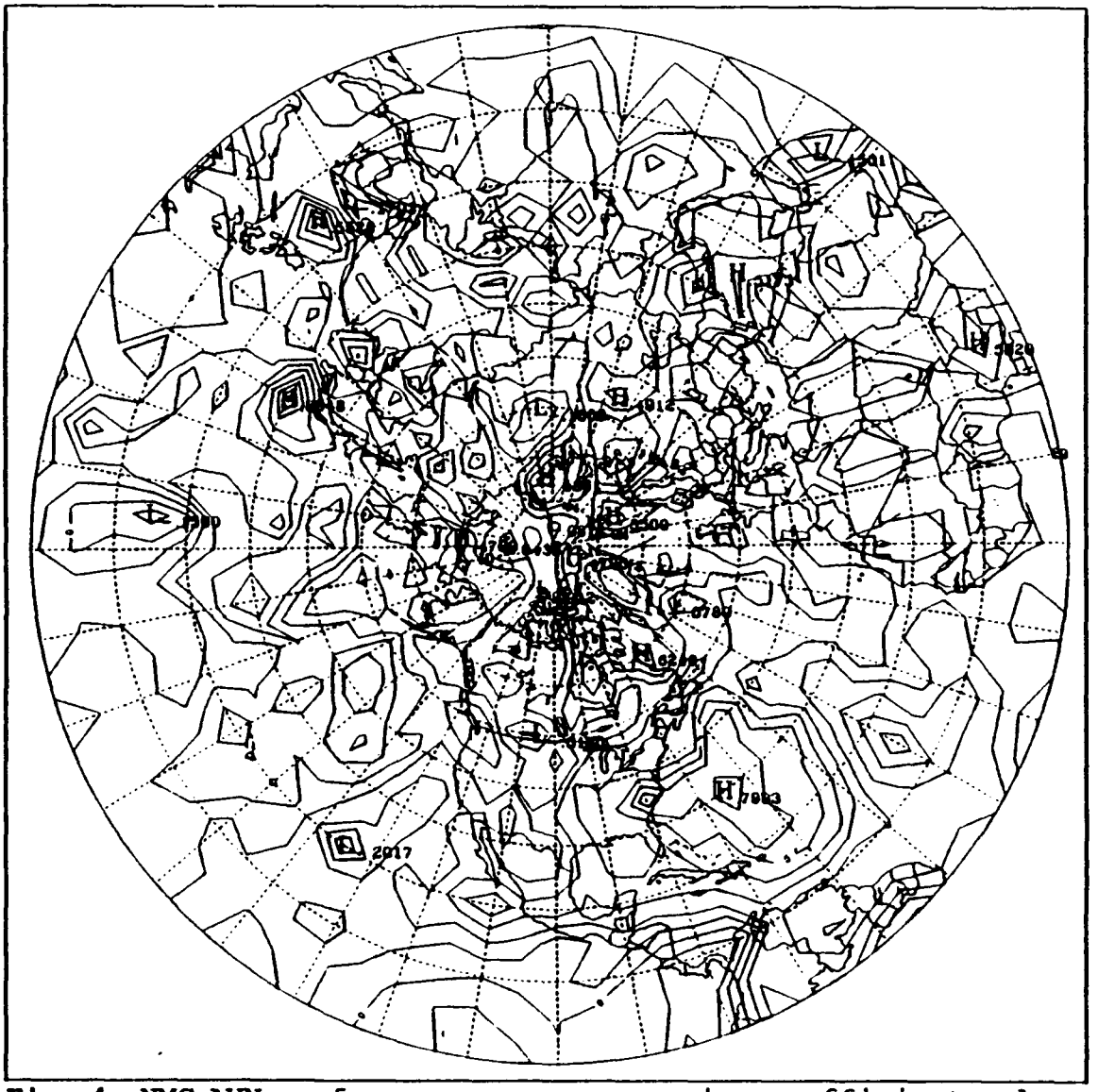

Fig. 4. NMC AVN surface pressure regression coefficient values (solid, contour interval 0.1) for the dependent data set (April-October 1992).

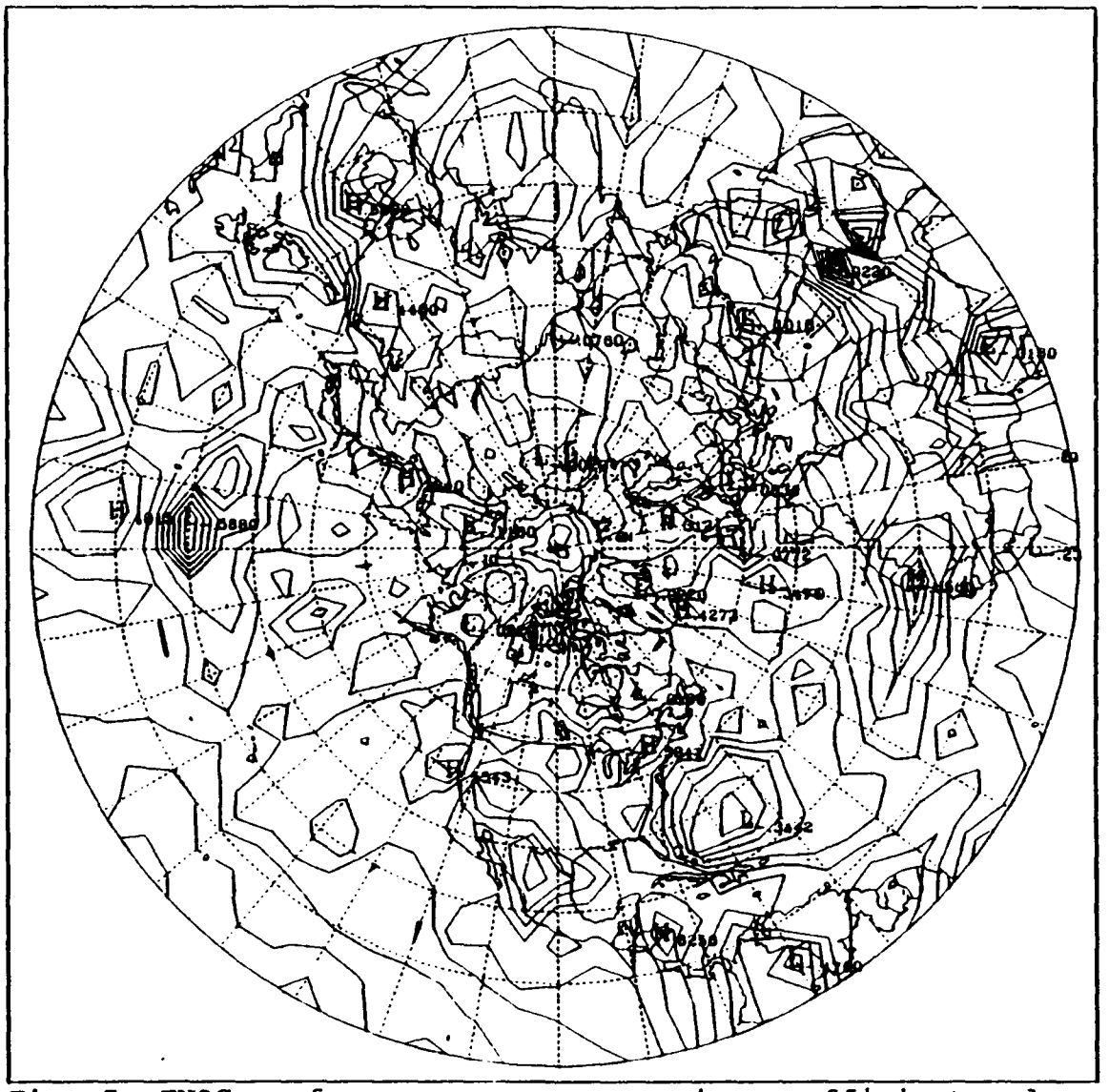

Fig. 5. FNOC surface pressure regression coefficient values (solid, contour interval 0.1) for the dependent data set (April-October 1992).

6) and corresponding 500-hPa intercept values (not shown) exhibited a more scattered distribution and established trends were less clear.

b. Ensemble Forecasts from Regression Method

Once the calculated regression coefficients and intercepts from the April through October dependent data set were obtained, the regression equations were then applied to independent forecast data to form ensemble forecasts and test the performance. The individual grid point regression equations for the 72-h forecast at both the surface and 500 hPa were applied to forecasts from the November data set. Due to a limited data set of extended range forecasts during November, the regression equations for the extended range forecasts were applied to the December data set.

Each application of the regression equations to a particular set of independent daily forecasts yielded ensemble forecasts for both the 72-h and extended range forecast periods at the surface and 500 hPa. The ensemble forecasts were then compared to the verifying analyses using a root mean-squared (RMS) error analysis method.

c. Smoothed Regression Ensemble Forecast

Consideration was given to the possibility that the ensemble forecasts produced by the independent application of the regression equations to each grid point might prove to be somewhat "noisy" or meteorologically unrealistic when

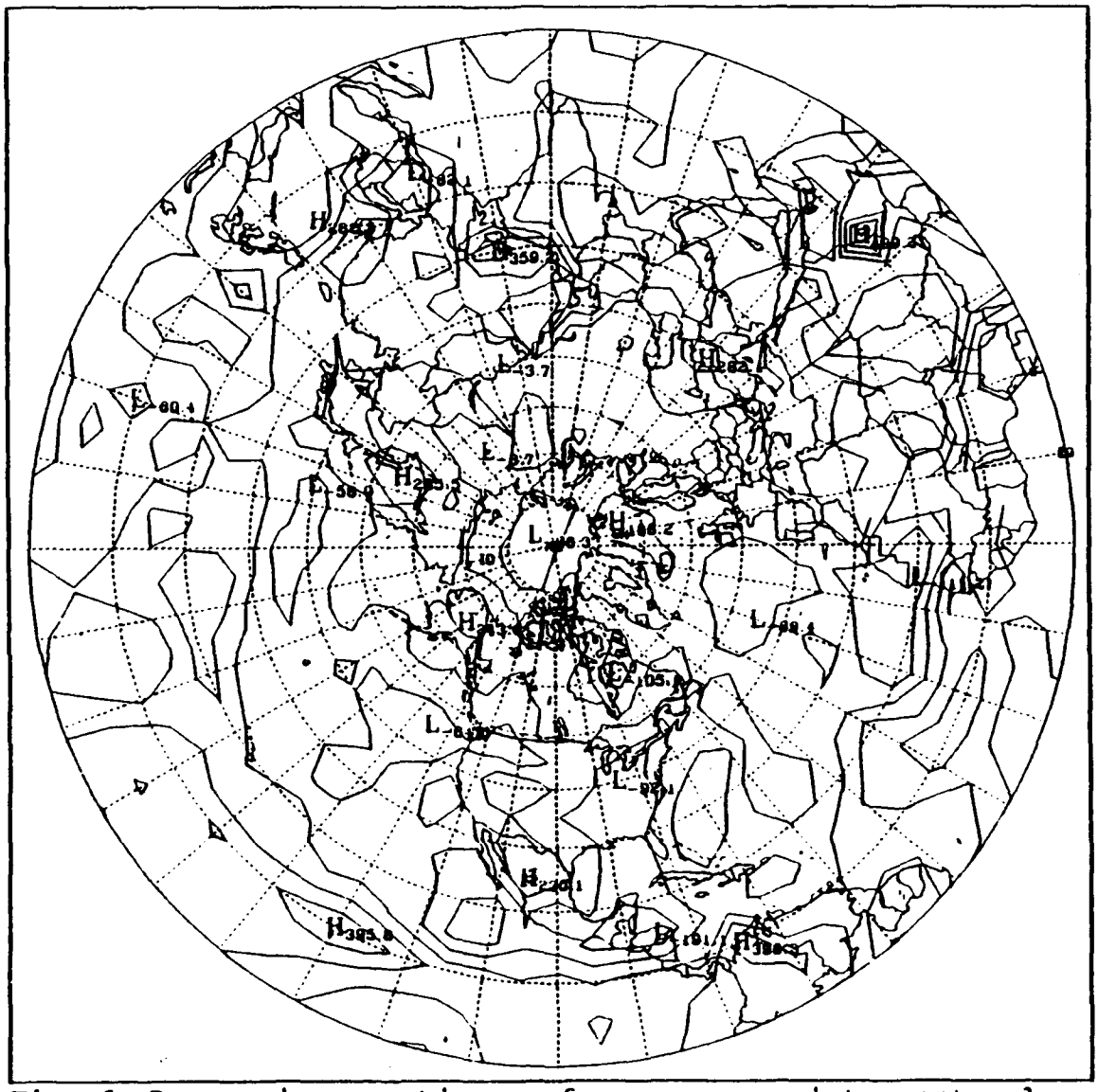

Fig. 6. Regression equation surface pressure intercept values (solid, contour interval 100 hPa) for the dependent data set (April-October 1992).

contoured. A smoothed version of the regressed ensemble forecast was therefore produced to address this potential concern. The smoothed regression forecast was developed by applying a limited five-point averaging of grid point data at each grid point in the final unsmoothed regression forecast. This was done by averaging the individual grid point values together with immediately adjacent north-south and east-west grid point values, totalling an average over five points.

4. Equally Weighted Forecast

As an added tool in determining the possible gain of the regressed ensemble forecast, an equally weighted ensemble forecast was produced for an independent data set corresponding with the set used in investigating the regressed forecasts. This technique approximates the model consensus forecast technique used by forecasters. The equally weighted forecast consisted of the compilation of an ensemble forecast from equal, one-third, contributions of the NMC, FNOC, and ECMWF products.

An equally weighted forecast was produced in every case where a regressed forecast was generated for the independent data set. The comparison of the equally weighted forecasts to the regressed forecasts was conducted to determine whether a significant gain in reduced forecast error would be obtained by optimizing forecast contributions through the regression method. Any additional gain by the regression

method would indicate in part that systematic error from the individual model forecasts was being reduced, which was not accounted for in the equally weighted forecasts.

5. RMS Error Analysis of Forecast Verification

In order to perform a baseline check of the effectiveness of the regression method in optimizing the ensemble forecast combination, a RMS error data analysis was first conducted for the dependent data set. The error analysis was then done to assess the ensemble forecast performance on the independent data sets.

The error analysis consisted of calculating the RMS error between the forecasts and the verifying analysis at each grid point in the domain. The RMS errors were then averaged over selected domains for comparison of forecast errors. The selected regions for the error analysis included the entire northern hemisphere, a region encompassing the northern hemisphere from 35°N to 90°, and a region extending from 50°N to 90°N. The error analysis was conducted for both the 72-h and extended range forecasts at both the surface and 500 hPa.

a. Dependent Data Set Analysis

The dependent data set consisted of the 72-h forecast and analysis grids from the April through October data period. The RMS error analysis was used in the verification of both ensemble and individual forecasts. The results of the analysis are presented in Table 3 and Table 4.

TABLE 3. RMS ERROR ANALYSIS OF 72-h SURFACE PRESSURE FORECAST VERIFICATION FOR THE DEPENDENT DATA SET (APRIL-OCTOBER 1992).

FORECAST TYPE/ VERIFYING ANALYSIS	RMS ERROR (mb) 0°-90°N	35°-90°N	50°-90°N
FNOC/FNOC	3.925	5.205	5.665
NMC AVN/AVN	3.615	4.544	4.920
ECMWF/ECMWF	3.316	4.378	4.785
1/3 CONTRIBUTION/ ECMWF	3.149	4.078	4.449
UNSMOOTHED REGRESSION/ECMWF	2.791	3.726	4.099
SMOOTHED REGRESSION/ECMWF	2.933	3.869	4.197

Within the dependent data set, the unsmoothed regression methods posted significant gains over both the individual and equally weighted forecasts. For example, the unsmoothed regression forecasts exhibited 1.479 hPa and 0.352 hPa error reductions when compared with the respective FNOC and equally weighted surface forecasts in the 35°-90°N region. At 500 hPa, the error reductions when compared to the FNOC and equally weighted forecasts were 13.863 m and 4.094 m, respectively, in the same region. The equally weighted forecasts also exhibited a reduction in mean forecast error over the individual forecasts in almost every case, indicating that even a raw ensemble forecast using equal weighting of the forecast inputs can reduce the error of the individual forecasts.

TABLE 4. RMS ERROR ANALYSIS OF 72-h 500-hPa HEIGHT FORECAST VERIFICATION FOR THE DEPENDENT DATA SET (APRIL-OCTOBER 1992).

FORECAST TYPE/ VERIFYING ANALYSIS	RMS ERROR (meters) 0°-90°N	35°-90°N	50°-90°N
FNOC/FNOC	36.720	51.351	56.943
NMC AVN/AVN	39.021	48.575	53.341
ECMWF/ECMWF	31.969	44.036	48.522
1/3 CONTRIBUTION/ ECMWF	32.486	41.582	45.554
UNSMOOTHED REGRESSION/ECMWF	28.266	37.488	41.474
SMOOTHED REGRESSION/ECMWF	28.721	39.029	42.821

b. Independent Data Set Analysis

The independent data set consisted of model data from November and December. This independent data set represented data from a fall or early winter meteorological regime. Use of this data was necessitated by the lack of an available independent data set corresponding with the summer regime of the dependent data set. The lack of seasonal correspondence in independent and dependent data sets was expected to place the regression method at a slight disadvantage when compared with its potential performance on an independent data set from a corresponding time period.

As previously mentioned, the November data set was utilized in the 72-h forecast study while a December data set was utilized in the extended range forecast study. Regressed

and equally weighted forecasts were produced from the independent forecast data sets and were verified along with the individual forecasts against the ECMWF analysis corresponding to the appropriate valid time of the forecasts. The unsmoothed and smooth versions of the regressed forecasts were included in this analysis.

An error analysis of the surface forecast verification for the 72-h forecasts during November is shown in Table 5. The unsmoothed regression forecasts provided the smallest RMS errors in all regions. The equally weighted and smoothed regression ensemble forecasts displayed competitive error values when compared with the unsmoothed regression forecasts, while the individual forecasts provided the highest error values. The unsmoothed regression forecasts showed the greatest error reduction over individual forecasts in the higher latitude regions. For example, the error at the surface was reduced by 1.3 and 2.1 hPa for the unsmoothed regression forecasts compared to the AVN and FNOC products, respectively.

Corresponding 72-h forecast error data at 500 hPa are displayed in Table 6. At 500 hPa, the three ensemble forecasts again had relatively similar error. The RMS error values of the equally weighted and unsmoothed regression forecasts in all regions varied less than one meter from each other. This suggests that on the independent data set there was little system error that was removed by the regression

TABLE 5. RMS ERROR ANALYSIS OF 72-h SURFACE PRESSURE FORECAST VERIFICATION FOR THE INDEPENDENT DATA SET (NOVEMBER 1992).

FORECAST TYPE	RMS ERROR (mb) 0°-90°N	35°-90°N	50°-90°N
FNOC	4.541	6.092	6.555
NMC AVN	4.013	5.371	5.774
ECMWF	3.594	4.782	5.126
1/3 CONTRIBUTION	3.351	4.398	4.736
UNSMOOTHED REGRESSION	3.171	4.226	4.498
SMOOTHED REGRESSION	3.267	4.349	4.603

technique at 500 hPa. The regressed forecasts, however, showed significant gains over the individual forecasts. The error reduction by using the ensemble forecasts was again larger for the higher latitude regions than the tropics, although the overall reduction was not as large as for the surface 72-h forecast data.

A similar analysis for the 108-h surface forecasts is summarized in Table 7. As expected, the error values for the extended range surface forecasts were substantially higher than those of the 72-h surface forecasts. The ensemble forecast methods again showed substantial error reduction over the individual forecasts. The unsmoothed regression forecasts also performed slightly better than the equally weighted forecasts. In general, the error reductions for the ensemble 108-h forecasts were larger than the error reductions for the

TABLE 6. RMS ERROR ANALYSIS OF 72-h 500-hPa HEIGHT FORECAST VERIFICATION FOR THE INDEPENDENT DATA SET (NOVEMBER 1992).

FORECAST TYPE	RMS ERROR (meters) 0°-90°N	35°-90°N	50°-90°N
FNOC	41.227	56.649	60.687
NMC AVN	44.201	55.885	59.976
ECMWF	34.812	47.298	50.083
1/3 CONTRIBUTION	33.088	42.852	45.712
UNSMOOTHED REGRESSION	32.954	43.345	46.512
SMOOTHED REGRESSION	33.302	44.770	48.051

ensemble 72-h forecasts relative to their respective individual forecasts. The one notable exception was the error reduction of the ensemble extended range forecasts relative to the NMC MRF products. The ensemble provided only a 0.1 hPa error reduction relative to the MRF forecast at the surface. This was surprising as the error reduction of the ensemble 72-h forecasts relative to the NMC AVN products was 1.3 hPa. The lack of error reduction relative to the MRF is difficult to explain and should be investigated more completely in a future study.

Verification data for the 500-hPa extended range forecasts are shown in Table 8. Overall, the smoothed regression products provided the lowest RMS errors in each region, however, the maximum difference between the RMS errors of the three ensemble forecast methods was less than one and

TABLE 7. RMS ERROR ANALYSIS OF EXTENDED-RANGE SURFACE PRESSURE FORECAST VERIFICATION FOR THE INDEPENDENT DATA SET (DECEMBER 1992).

FORECAST TYPE	RMS ERROR (mb) 0°-90°N	35°-90°N	50°-90°N
FNOC	7.781	11.090	12.552
NMC MRF	5.935	8.214	8.746
ECMWF	6.322	8.831	9.818
1/3 CONTRIBUTION	5.679	7.885	8.748
UNSMOOTHED REGRESSION	5.588	7.803	8.621
SMOOTHED REGRESSION	5.583	7.800	8.617

one-half meters for all regions. As with the surface 108-h forecasts, the errors for the 500-hPa 108-h forecasts were significantly higher than those of the 72-h 500-hPa forecasts. Although little variation existed in the RMS errors of the ensemble forecasts, the ensemble forecast errors represented a substantial reduction in RMS error when compared with the errors of the individual forecasts. Additionally, there was a general trend of increased error reduction in ensemble forecasts relative to the individual forecasts for higher latitude regions. For example, the unsmoothed regression forecasts showed an error reduction of 30.008 m relative to the individual FNOC forecasts in the high latitude regions while the error reduction was only 16.029 m when computed over the entire hemisphere.

TABLE 8. RMS ERROR ANALYSIS OF EXTENDED-RANGE 500-hPa HEIGHT FORECAST VERIFICATION FOR THE INDEPENDENT DATA SET (DECEMBER 1992).

FORECAST TYPE	RMS ERROR (meters) 0°-90°N	35°-90°N	50°-90°N
FNOC	69.861	98.286	105.413
NMC MRF	63.999	83.422	87.176
ECMWF	57.324	79.268	85.232
1/3 CONTRIBUTION	53.693	71.688	76.224
UNSMOOTHED REGRESSION	53.832	71.360	75.405
SMOOTHED REGRESSION	52.370	70.833	74.786

c. Discussion

As would be expected, the error reduction of the unsmoothed regression forecasts over the remaining individual and ensemble forecasts was more pronounced in the dependent data set than in the independent data sets. This occurred primarily because the regression equations were extracted from the dependent data. However, the different seasonal nature of the two data sets was likely a secondary factor in narrowing the gain of the unsmoothed regression method over the remaining forecast methods. Although the unsmoothed regression method retained superiority over the individual forecasts in the independent data set, the margin of difference in error reduction between the regression methods and the equally weighted forecasts was not as notable. Use of

the regression method over an independent data set from a summer meteorological regime would likely show greater error reduction than found for the November and December data. Success of the regressed forecasts is tied to the ability of the regression method to identify systematic error in the individual models and reduce this error by optimizing individual model contributions. Systematic errors identified in the summer regime may not apply in all cases to the fall or winter regime.

Although the possibility existed that plots of the unsmoothed regression forecasts might appear "noisy" when contoured, the resulting products appeared meteorologically sound. In fact, the smoothing conducted to obtain the smoothed regression products proved to be unnecessary. In terms of error statistics, the smoothed regression products were slightly inferior to the unsmoothed regression products within the dependent data set and to the 72-h forecasts within the independent data set. The smoothed regression products slightly outperformed the unsmoothed products for the extended range forecasts in the independent data set, however, these improved results did not appear to be linked to any meteorological explanation. Because the differences between the smoothed and unsmoothed regression forecasts were small, only the performance of the unsmoothed regression forecasts are considered for the case studies in the next section.

IV. CASE STUDIES

A. INTRODUCTION

The RMS error analysis of the previous sections indicated the overall performance trends of the various forecast models. However, the RMS error values did not provide a good representation of the day-to-day differences that occur in the individual forecasts or the success with which the ensemble forecasts handled these differences. Poor performances of daily forecasts are smoothed or averaged in the computation of RMS errors. In order to determine the magnitude or characteristics of the improvement for individual forecasts achieved through the ensemble forecasting method, individual case studies were conducted.

The following case studies were selected to represent events where significant variation existed in individual forecasts that covered the geographical region from the eastern Pacific Ocean to the eastern United States from 25°-75°N. Within this region, the ECMWF, NMC, and FNOC analyses were subjectively compared to identify any significant differences. The individual forecasts of selected synoptic features were then compared subjectively to note differences in model forecast values. The ensemble forecasts of these features were then subjectively compared to the individual

model forecast features to determine how individual model forecast differences might have affected the ensemble combinations. Finally, a verification of both individual and ensemble forecasts was conducted using the ECMWF analyses as the verification analyses.

B. CASE 1 (1200 UTC 8 AUGUST, 1992)

1. Comparison of Surface Pressure Analyses

The primary surface features on the FNOC (Fig. 7a), NMC (Fig. 7b), and ECMWF (Fig. 7c) analyses of sea-level pressure were the central Canadian low pressure area, the thermal trough in the southwestern United States, the Gulf of Alaska low, and the eastern Pacific Ocean high pressure area. The central Canadian low on the ECMWF analysis was characterized by a broad low pressure area and south to southeastward extending trough patterns. Similar positions and structures of the low were indicated on the NMC and FNOC analyses. Central pressures for the low in the three analyses were comparable while a maximum difference of 1.3 hPa existed between the FNOC and ECMWF analyses. The thermal trough in the southwestern United States was characterized by a weak pressure gradient, and was represented as weak troughing on the ECMWF analyses. The NMC analysis indicated a weak closed low in the area while the FNOC analysis indicated weak troughing in a different orientation compared to the ECMWF analysis. However, the maximum pressure differences between

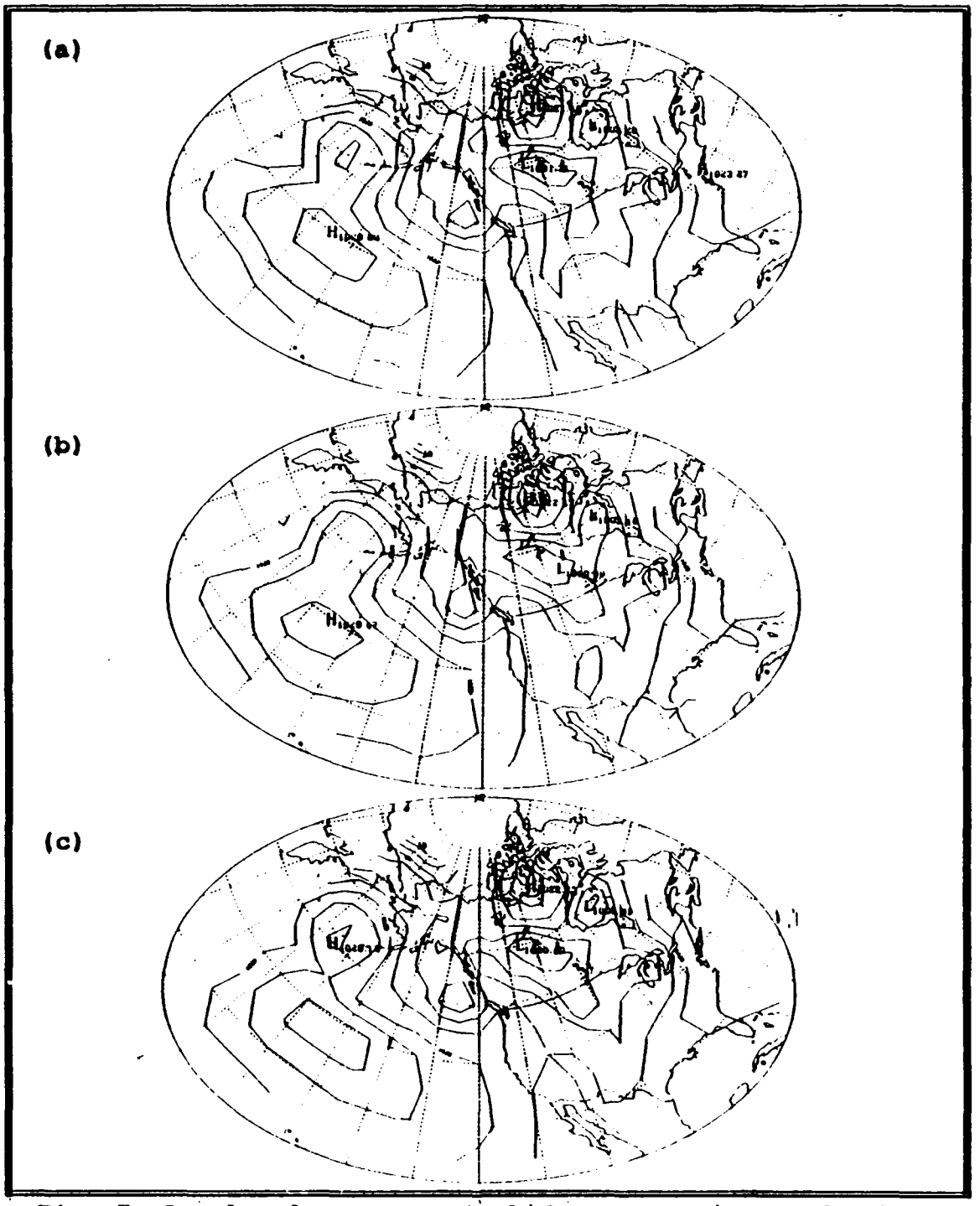

Fig. 7. Sea-level pressure (solid, contour interval 4 hPa) analyses at 1200 UTC 8 August 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

the three analyses were generally less than approximately 2 hPa. The Gulf of Alaska low was represented as a closed low on the ECMWF analysis with central pressures less than 1008 hPa. The low was represented very similarly in pressure and position on the NMC and FNOC analyses. The eastern Pacific high was characterized by a high pressure ridge with a lobe of ridging over the Bering Sea on the ECMWF analysis. Similar representations of the high and the associated ridging were exhibited by the NMC and FNOC analyses. However, the ECMWF and FNOC analyses captured a relative maximum surface pressure center in the lobe of ridging over the Bering Sea which was missed by the NMC analysis.

Based on these comparisons of individual synoptic features, the use of the ECMWF analysis for verification in this case study should not produce any major errors relative to any other analysis.

2. Surface Pressure Forecast Comparison and Verification

A comparison of the individual (Fig. 8a-c) and ensemble (Fig. 9a,b) forecasts through difference plots (Fig. 10a-c and Fig. 11a,b) revealed that the regression forecast provided the lowest overall errors in forecasting a majority of the main synoptic features at the surface for this case study. Details of the forecast comparisons and verification are provided in the following discussion of synoptic features.

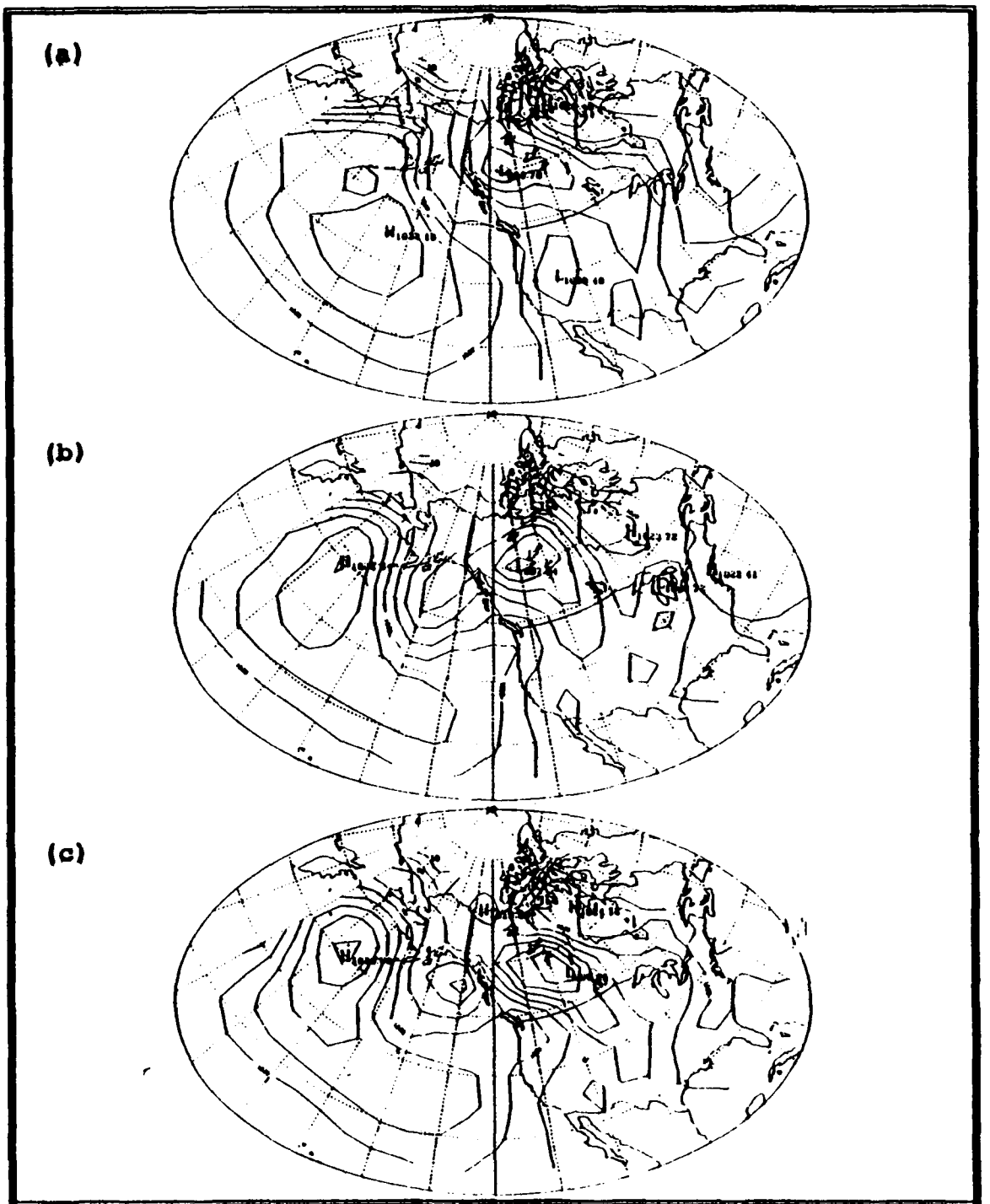

Fig. 8. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 8 August 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

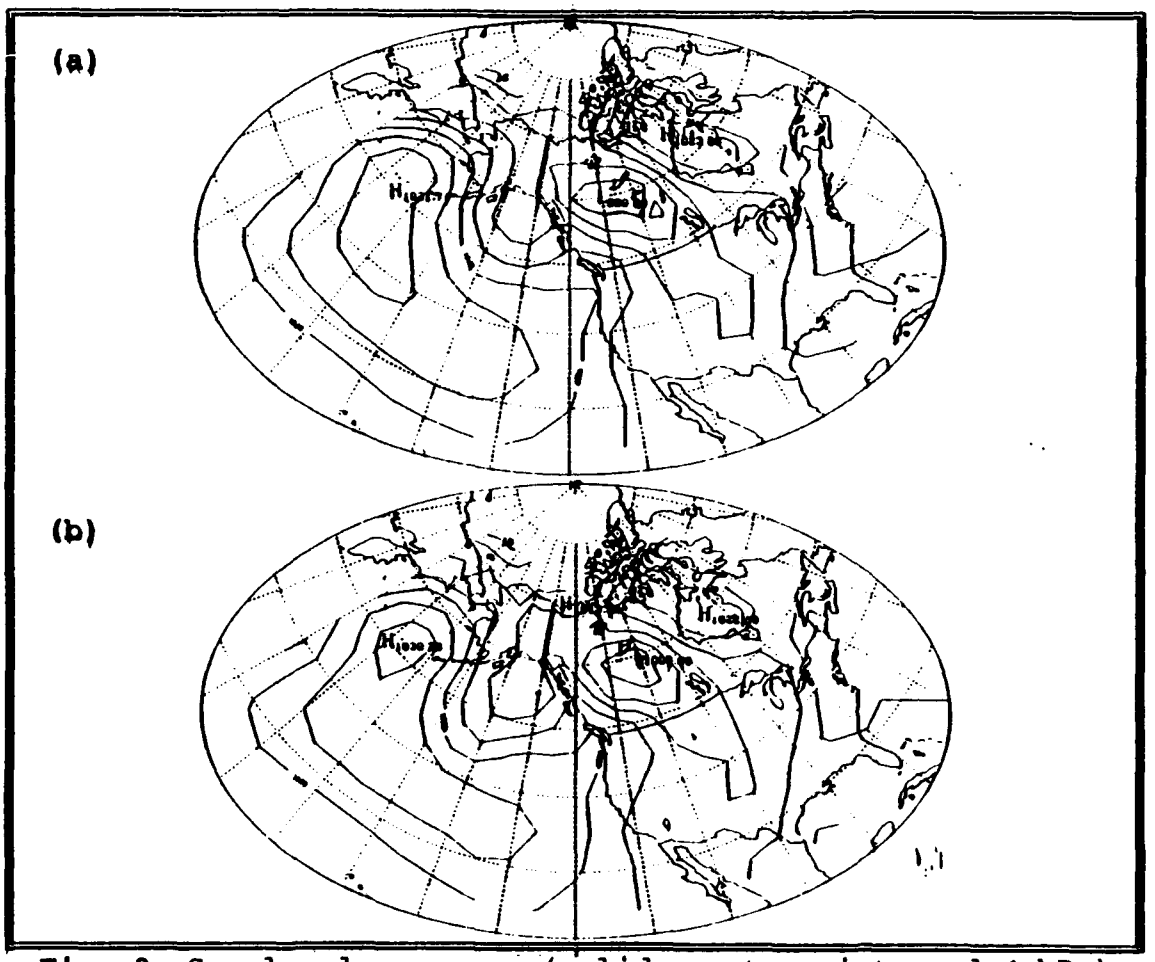

Fig. 9. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 8 August 1992 from the (a) equally weighted and (b) regression ensembles.

a. Canadian Low and Associated Trough Pattern

The Canadian low pressure center was the dominant feature over North America. The FNOC (Fig. 8a), AVN (Fig. 8b), and ECMWF (Fig. 8c) forecasts placed this low over Canada but their positions differed substantially and all three were too intense as indicated by the FNOC (Fig. 10a), AVN (Fig. 10b), and ECMWF (Fig. 10c) error plots. The difference in the positions of the low on the individual model forecasts is the result of phase differences between the forecasts. For example, the relative difference in the forecast position of the low in the FNOC and ECMWF forecasts was approximately 20° of longitude. The equally weighted and regressed forecasts placed the low at an intermediate position corresponding to an overall blend of the individual forecast positions. As shown in the error plots for the equally weighted (Fig. 11a) and regression (Fig. 11b) forecasts, the overall error in the region was reduced by the ensemble forecasts. The regression forecast placed the low at a more accurate position than the individual models with the most accurate intensity.

The orientation of the southward extending trough varied significantly between the individual forecasts due mainly to phase differences between the models. The AVN forecast trough axis (Fig. 8b) was approximately 10° of longitude west of the FNOC trough axis (Fig. 8a), while the

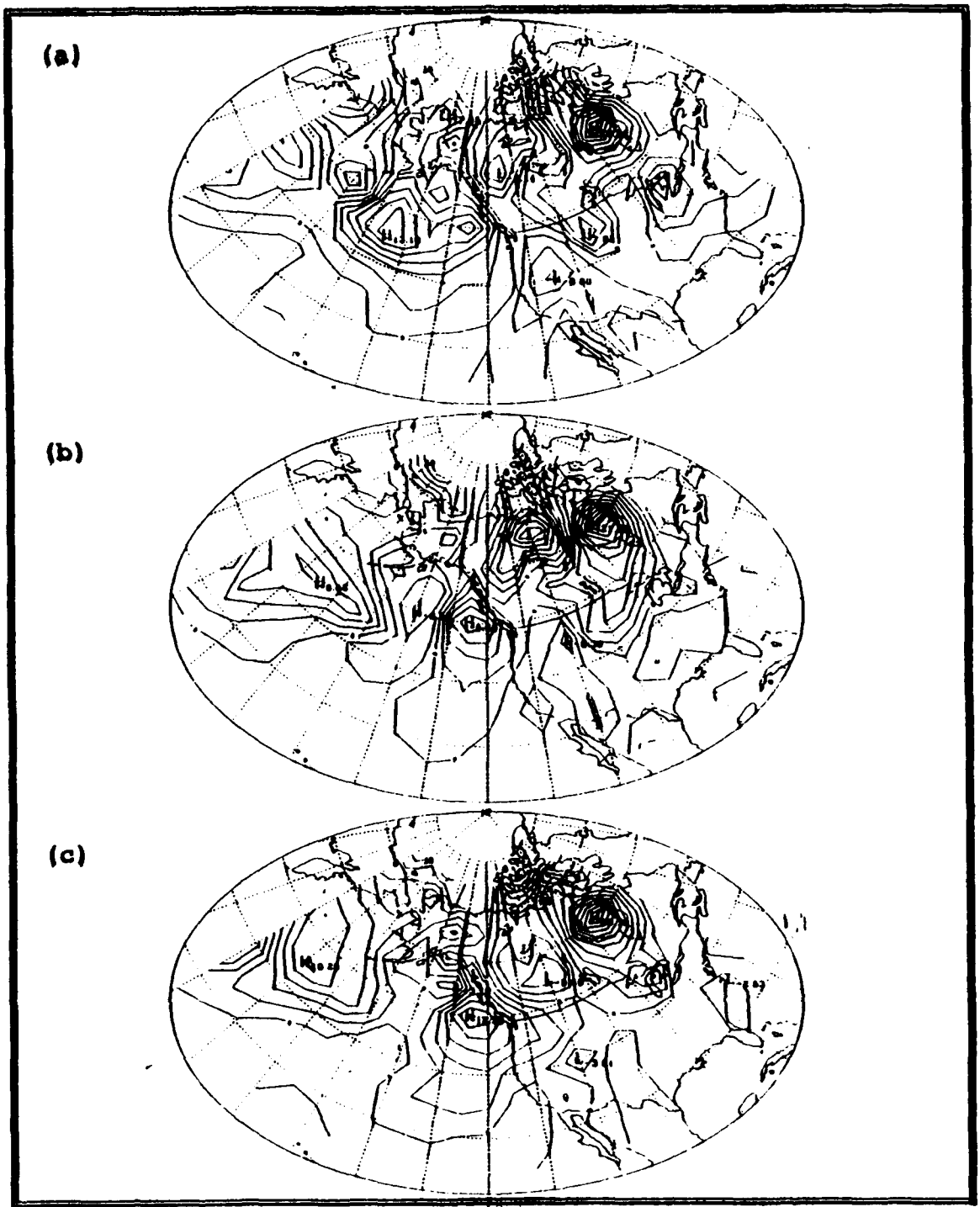

Fig. 10. Difference plots of the 72-h (a) FNOC, (b) NMC AVN, and (c) ECMWF sea-level pressure forecasts verified with the ECMWF analysis at 1200 UTC 8 August 1992 (2 hPa contours).

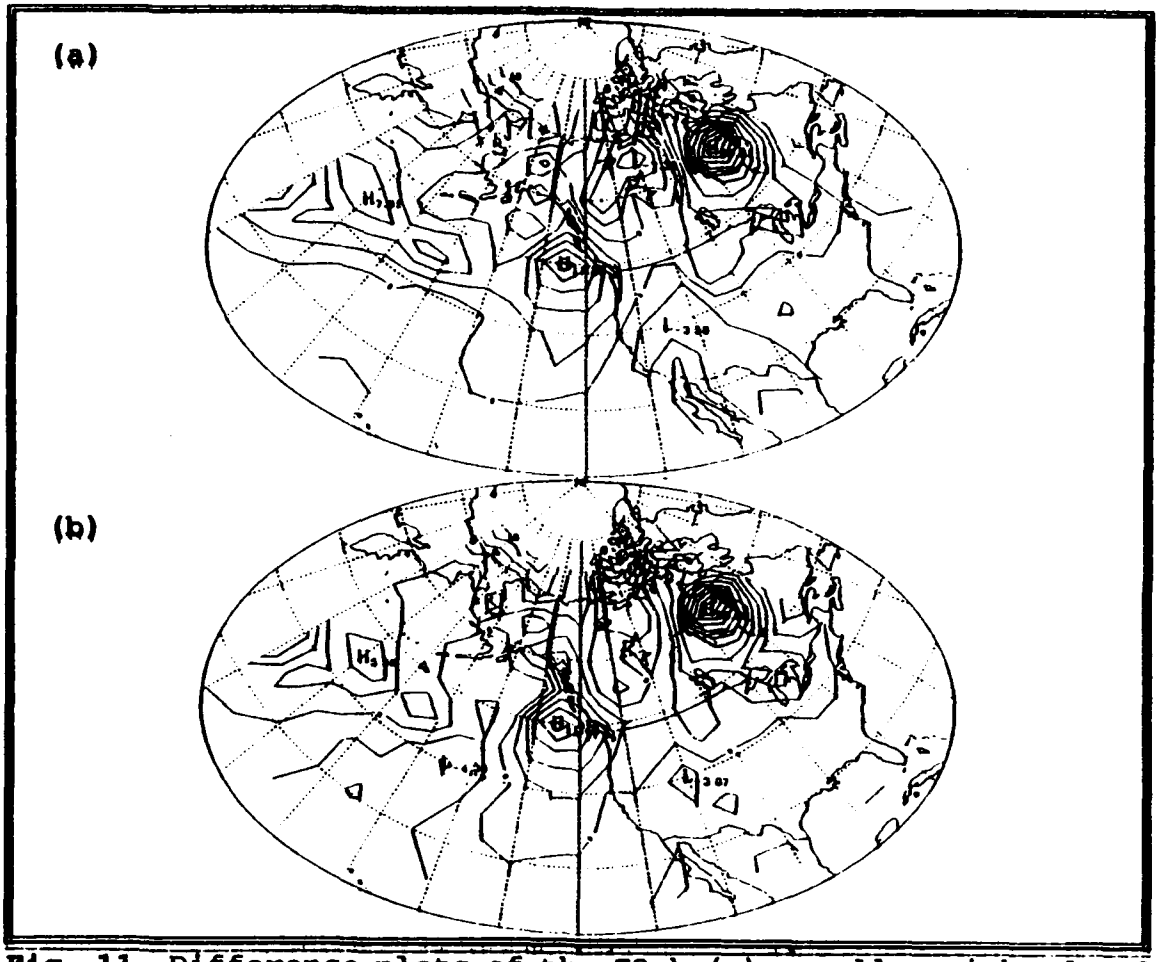

Fig. 11. Difference plots of the 72-h (a) equally weighted and (b) regression sea-level pressure forecasts verified with the ECMWF analysis at 1200 UTC 8 August 1992 (2 hPa contours).

ECMWF axis (Fig. 8c) was forecast at an intermediate position. The ECMWF forecast produced a continuous extension of the trough into the south central United States, while the AVN and FNOC products forecast disconnected low pressure areas over the south central United States which were not present on the ECMWF analysis. The regressed and equally weighted forecasts (Fig. 9a,b) corresponded most closely in trough orientation to the intermediate position of the ECMWF forecast and coincided with the ECMWF analysis. The regression forecast provided the most accurate position and intensity of the trough feature as indicated by the low error values in the region (Fig. 11b). These error values were generally less than 2 hPa over the region.

b. Thermal Low in Southwestern United States

The regressed and individual forecasts exhibited a thermal low pressure area in the southwestern United States, however, the areal extent of the low differed in the forecasts. The regressed forecast (Fig. 9b) exhibited a slack gradient which was more in line with the ECMWF analysis and verified most accurately with the lowest overall error in the region (Fig. 11b). The individual model forecasts exhibited the stronger gradients in the area, in general, and the FNOC forecast (Fig. 8a) exhibited the largest area of low pressure less than 1012 hPa. The equally weighted forecast (Fig. 9a) exhibited no indication of a relative low in the area which

verified more accurately than the individual model forecasts, given the slack gradient in the region. However, the equally weighted forecast missed the structure evident in the regression forecast.

c. Great Lakes Low Pressure Area

This case represents an example where one forecast member may represent a weak feature such as the Great Lakes low well, however, its amplitude may not be sufficient to compensate for the missed forecasts from other models. In this region, only the AVN forecast (Fig. 8b) exhibited a relative minimum low pressure center which verified most accurately in terms of pressure in the immediate vicinity of the low. The ensemble forecasts did not perform as well for the feature due to poor forecast contributions from the FNOC and ECMWF forecasts which basically missed the feature. However, the regressed forecast did perform better than the ECMWF and FNOC forecasts in the area.

d. Gulf of Alaska Low

The comparison of forecasts for this feature provides an example of consistently poor forecasts by all three individual model forecasts. In reviewing the individual forecasts, the AVN product (Fig. 8b) forecast the deepest low pressure over a broad area in the region but did not represent a defined cutoff low as analyzed on the ECMWF analysis. The ECMWF forecast (Fig. 8c) exhibited a central pressure

comparable to the ECMWF analysis pressure for the low but forecast the low too far north. The FNOC product (Fig. 8a) forecast surface pressure in the area at least 4 hPa higher than the ECMWF analysis and did not forecast a low pressure center. The equally weighted and regressed forecasts (Fig. 9a,b) exhibited pressure values approximately 1-3 hPa higher than the ECMWF analysis and only the regressed forecast showed a low pressure center in the area. When verified, the equally weighted and regressed forecasts performed the best in the area while the AVN forecast provided slightly lower errors in the immediate vicinity of the low. However, all of the forecasts, including the regression forecast, poorly forecast this low. The lack of any ensemble member providing a correct forecast prevented the ensemble from providing much real improvement for this feature.

e. Eastern Pacific Subtropical High

The key element in this region is the high pressure area over the Bering Sea. The individual models missed the forecast of this feature in a variety of ways in terms of intensity and position. The random nature of these misses contributed to a good ensemble forecast through a blending of the individual forecasts.

Comparison of the specific forecast details of this feature revealed that the orientation of the high pressure region in the AVN and ECMWF forecasts (Fig. 8b,c) was similar

to that of the ECMWF analysis. However, the central pressure values near the Bering Sea in the AVN and ECMWF forecasts were approximately 3-7 hPa higher than the ECMWF analysis pressure. The orientation of the FNOC high pressure region, however, differed significantly from the ECMWF analysis and central pressure values were approximately 4 hPa lower than the analysis. This differing orientation of the FNOC forecast was largely influenced by the associated absence of the deeper low pressure region in the Gulf of Alaska represented in the other individual forecasts. The orientation and central pressures of the high pressure region in the regressed and equally weighted forecasts (Fig. 9a,b) were very similar to the ECMWF analysis.

As mentioned, the random nature of errors within the individual forecast contributions allowed the regressed forecast to verify most accurately near the high pressure area over the Bering Sea. In the eastern portion of the high pressure area or ridging to the east, several products were comparable. The regressed and equally weighted forecasts, however, appeared to show the best performance in terms of low error and weaker error gradients. The slight gains by the ensemble forecasts in the eastern area were a result of a greater correspondence in individual forecast contributions for this area.

3. Comparison of 500-hPa Height Analyses

The FNOC (Fig. 12a), NMC (Fig. 12b) and ECMWF (Fig. 12c) analyses had three major synoptic features which matched closely in most regions. These features were a closed upper-level low in the southeastern Gulf of Alaska and an associated major trough along the west coast of North America, a trough across the Great Lakes region, and ridging over the Bering Sea.

The Gulf of Alaska low and the southward extending trough were analyzed similarly by each center with some small differences noted in the height at the low center. The central height value of the NMC Gulf of Alaska low was analyzed approximately 9 m deeper than the corresponding central height in the FNOC product and 18 m deeper than the ECMWF height value.

The greatest analysis differences occurred with the weak trough over the Great Lakes. The NMC trough in the Great Lakes region was analyzed approximately 30 m deeper than the troughs in the ECMWF and FNOC products although their phasing was very similar.

Ridging over the Bering Sea was the other major synoptic feature in the analyses. This feature was quite similar in amplitude and position in all of the analyses.

Given the general similarities in the analyses, the use of the ECMWF analysis would be appropriate as the verification analysis in most areas. However, differences in

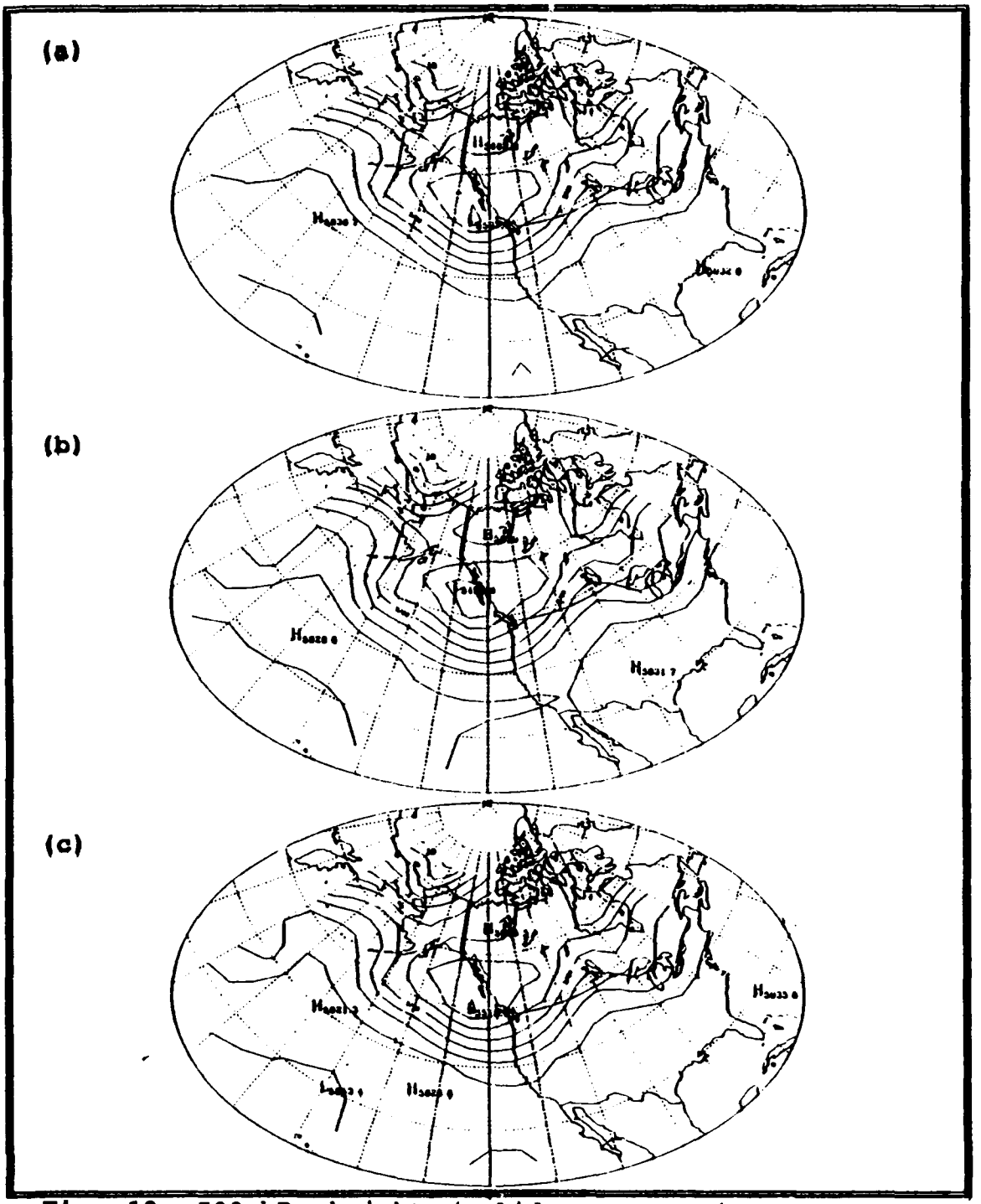

Fig. 12. 500-hPa height (solid, contour interval 60 m) analyses at 1200 UTC 8 August 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

the Great Lakes region could cause verification values to be suspect if forecast error values are comparable or smaller in magnitude than these analysis differences.

4. 500-hPa Height Forecast Comparison and Verification

Comparison of individual (Fig. 13a-c) and ensemble (Fig. 14a,b) 500-hPa height forecasts through difference plots (Fig. 15a-c and Fig. 16a,b) revealed mixed results in the performance of the models for this case study. The regression forecast provided the best results in the Great Lakes trough and Bering Sea ridging areas while the FNOC forecast performed the best in the region of the low near the Gulf of Alaska. Details of the forecast comparisons and verification are provided in the following discussion.

a. Great Lakes Trough

In this area, the random nature of the errors, mainly phase differences, in the individual model forecasts allowed the blend of the forecasts in the regression method to provide the best performance. Details of the forecast differences and the ensemble blending can be seen in the 500-hPa forecasts for 8 August which are provided in Fig. 13a-c and Fig. 14a,b. As shown in Fig. 13b, the AVN forecast trough was significantly deeper than the corresponding troughs in the ECMWF and FNOC forecasts (Fig. 13a,c) and the ECMWF analysis. Additionally, the FNOC trough position greatly lagged the trough position of the ECMWF analysis. The trough positions

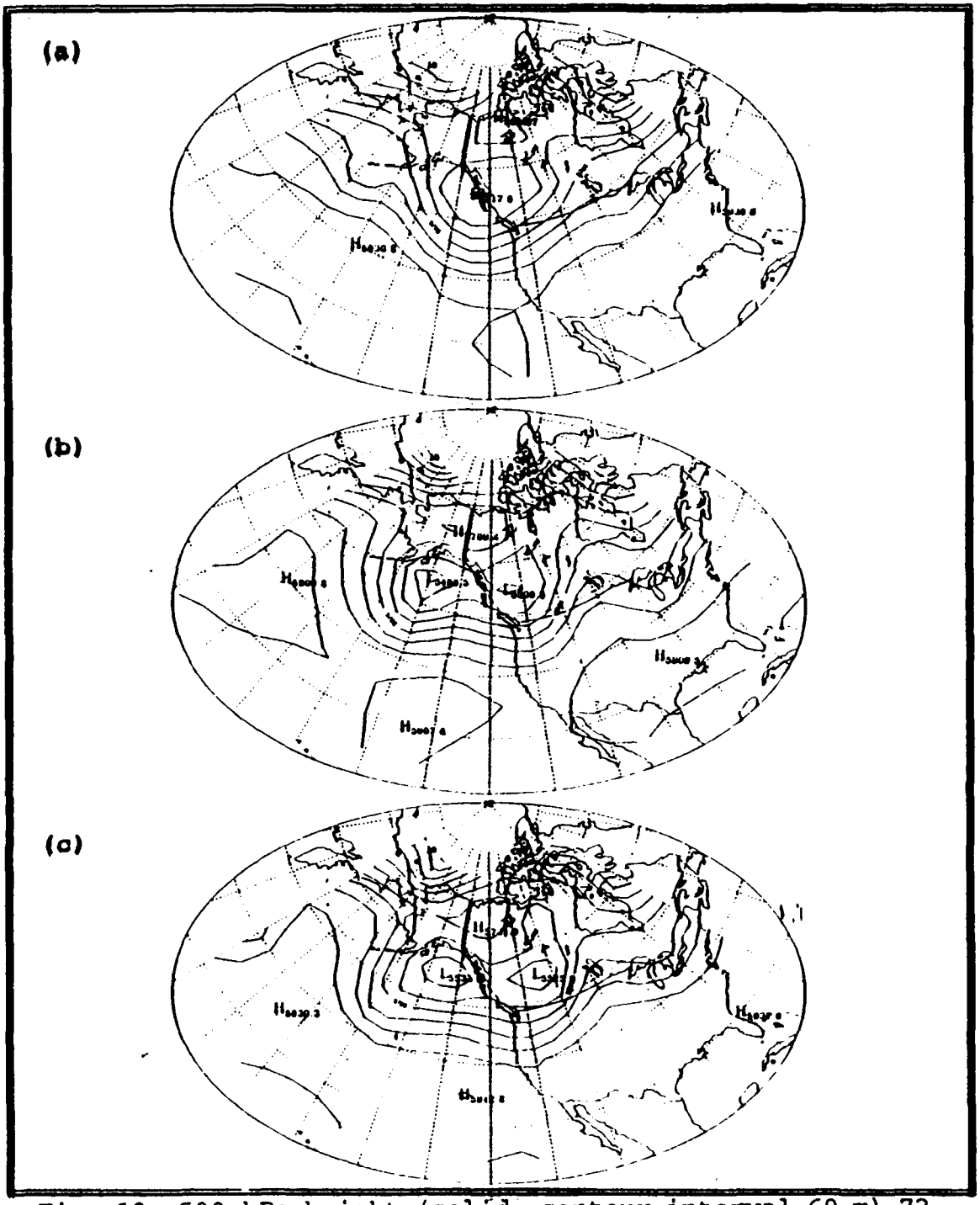

Fig. 13. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 8 August 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

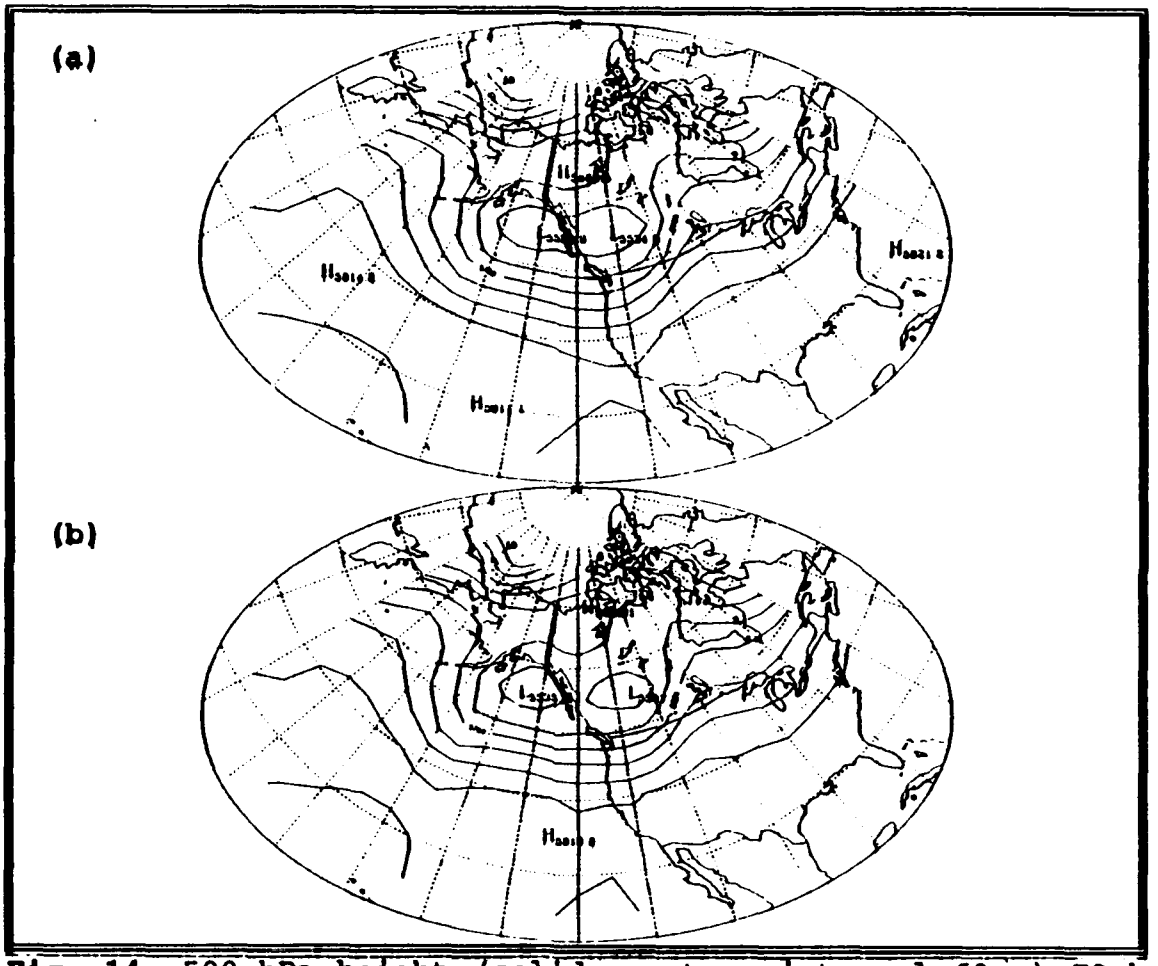

Fig. 14. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 8 August 1992 from the (a) equally weighted and (b) regression ensembles.

in the equally weighted and regressed forecasts (Fig. 14a,b) generally matched the trough position in the ECMWF analysis. The trough position and intermediate height values of the ensemble forecasts reflected a blended intermediate position and strength between the respective values of the individual forecasts. The blend in the regression method was an improvement over the individual forecasts as revealed by the difference plots. The error plot for the 500-hPa regression forecast (Fig. 16b) shows that the regression forecast had the lowest error for this feature.

b. Gulf of Alaska/Western Canada Low Pressure Area

This case exhibited an example where two of the individual forecasts, the ECMWF and AVN forecasts, performed poorly while the remaining FNOC forecast performed very well. The ensemble blend of the regression forecast was not able to compensate for the poor forecast contributions and was therefore outperformed by the FNOC forecast.

The AVN and ECMWF forecasts (Fig. 13b,c) provide an example of similar but inaccurate dual low scenarios. The AVN low positions in this feature lagged the ECMWF positions by 5°-10° of longitude. The FNOC product (Fig. 13a) forecast a single low center at a more accurate intermediate position between the forecast lows of the ECMWF and AVN products. The center height values in both ECMWF lows were forecast higher

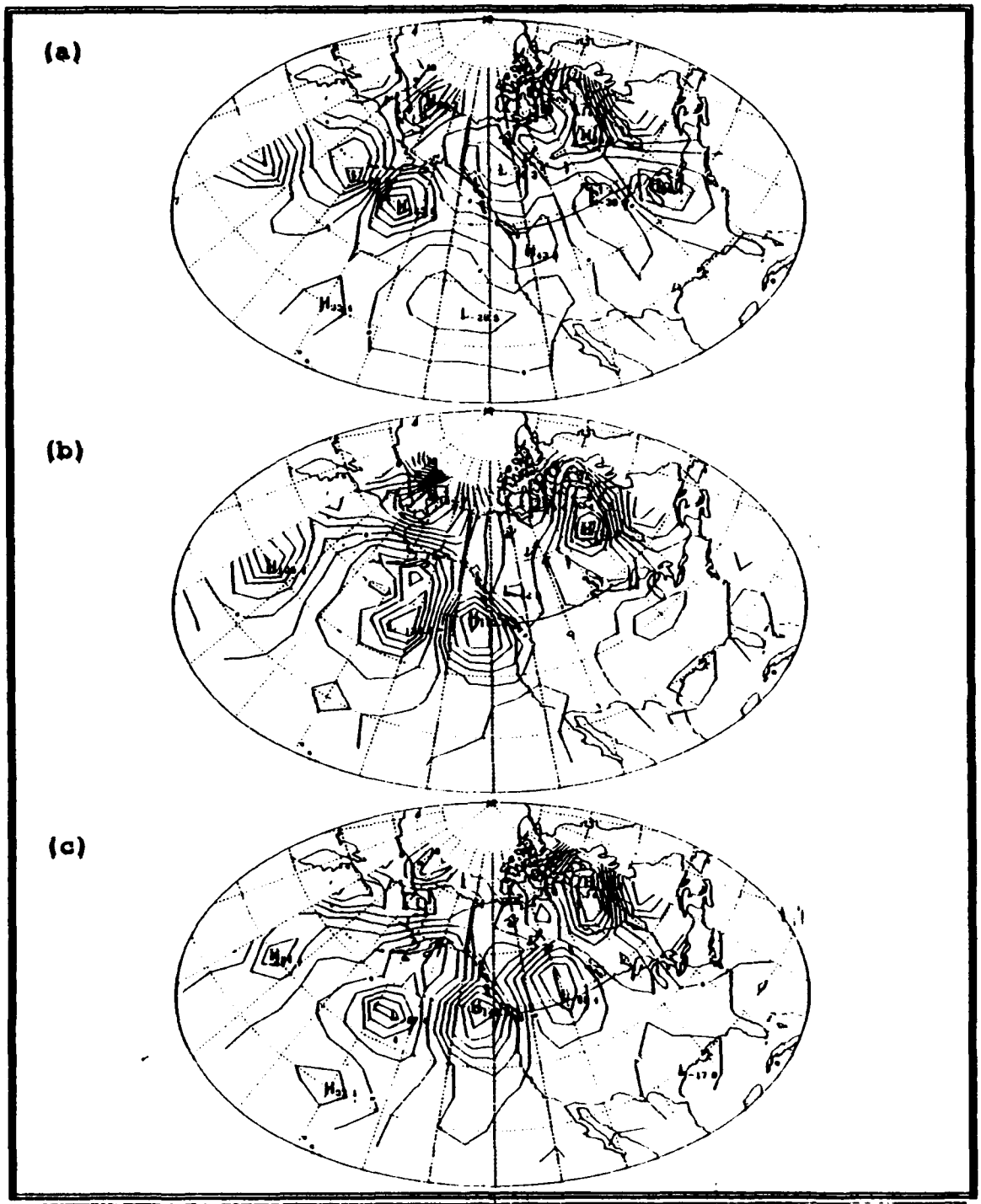

Fig. 15. Difference plots of the 72-h (a) FNOC, (b) NMC AVN, and (c) ECMWF 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 8 August 1992 (20 m contours).

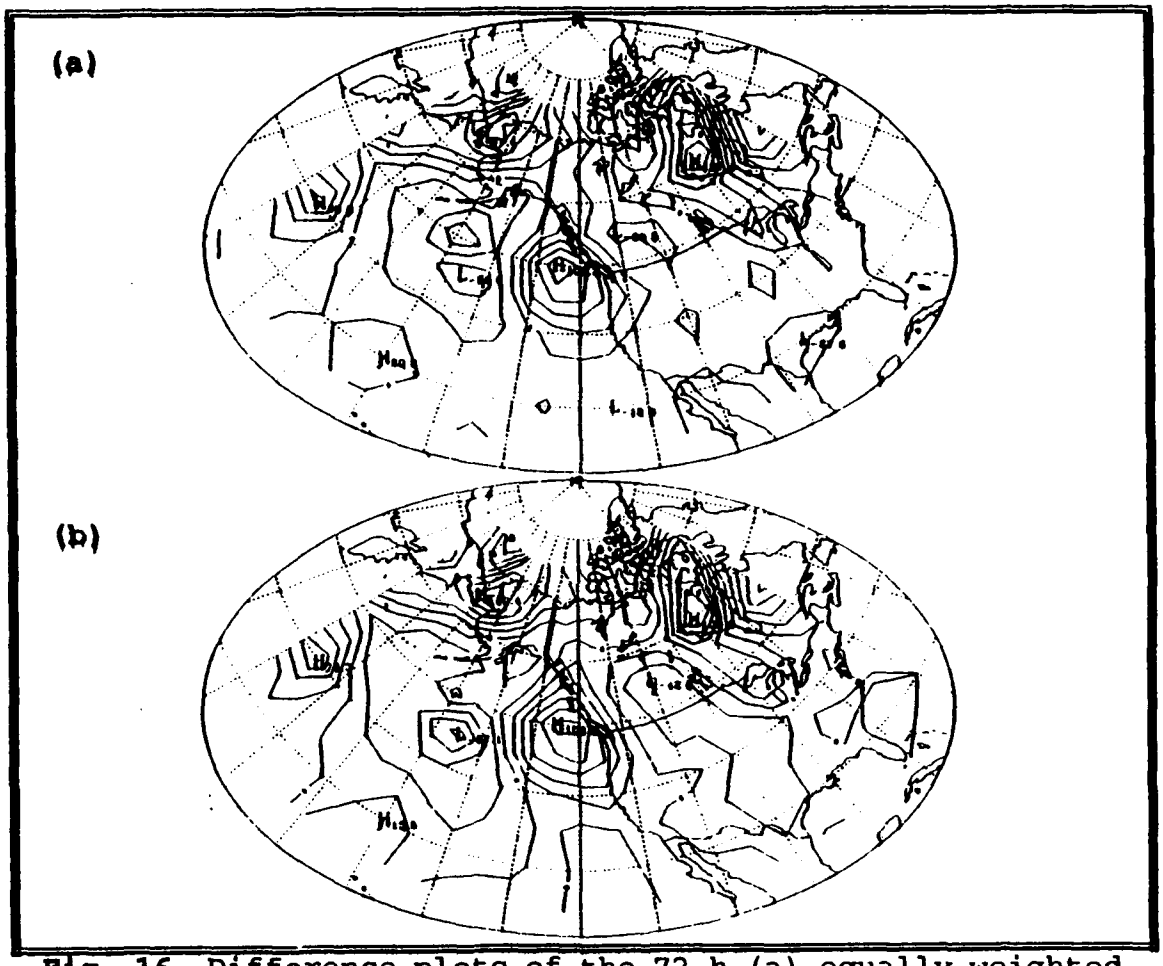

Fig. 16. Difference plots of the 72-h (a) equally weighted and (b) regression 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 8 August 1992 (20 m contours).

than the respective AVN lows, up to approximately 44 m for the western low. The intensity of the FNOC low was correctly forecast at an intermediate value between the ECMWF and AVN forecast values. The regressed and equally weighted forecasts (Fig. 14a,b) carried the dual low pattern with central heights similar to the ECMWF values.

When verified, the FNOC forecast of the single low pattern performed significantly better than all of the remaining forecasts. The forecast ridge between the lows of the dual low pattern was a large source of error for forecasts with this pattern. The equally weighted and regressed forecasts slightly outperformed the ECMWF and AVN forecasts, demonstrating a gain through the incorporation of the FNOC forecast through blending.

c. Bering Sea Ridging

In this area, the forecast differences of the individual model forecasts were somewhat random in position and intensity which allowed the ensemble blend of the regression method to provide the best performance in the area. Phase differences in the ridge axis were present in the individual forecasts in the upper portion of the ridge. In the lower portion of the area, the FNOC forecast inaccurately forecast a trough feature which was not present in the other two forecasts. When verified, the regressed forecast provided the best position and intensity of the ridge pattern.

C. CASE 2 (1200 UTC 7 SEPTEMBER, 1992)

The major synoptic differences in Case 2 when compared with Case 1 are a developing lee cyclone over the central United States and the existence of a well-defined low pressure area over the Bering Sea. Investigation of the lee cyclone feature, in particular, provides a feature which is more likely to have characteristic systematic errors in the individual model forecasts. Identification of this error by the regression method should provide forecast gains.

1. Comparison of Surface Pressure Analyses

The major features in the FNOC (Fig. 17a), NMC (Fig. 17b), and ECMWF (Fig. 17c) analyses of sea-level pressure are the Bering Sea low pressure area, the Gulf of Alaska coastal low, the central United States lee cyclone, and high pressure areas over the northwestern United States and the eastern Pacific Ocean. The positions of the high and low pressure centers generally corresponded well in the FNOC, NMC, and ECMWF analyses. A few amplitude differences did, however, exist in the analyses. For example, the NMC surface pressure value for the low in the Bering Sea was approximately 1-1.5 hPa higher than corresponding pressures on the other analyses. The NMC pressure value for the low over the central United States was also higher than the other products, approximately 2.7 hPa higher than the corresponding pressure on the ECMWF analysis. In the northwestern United States, the ECMWF

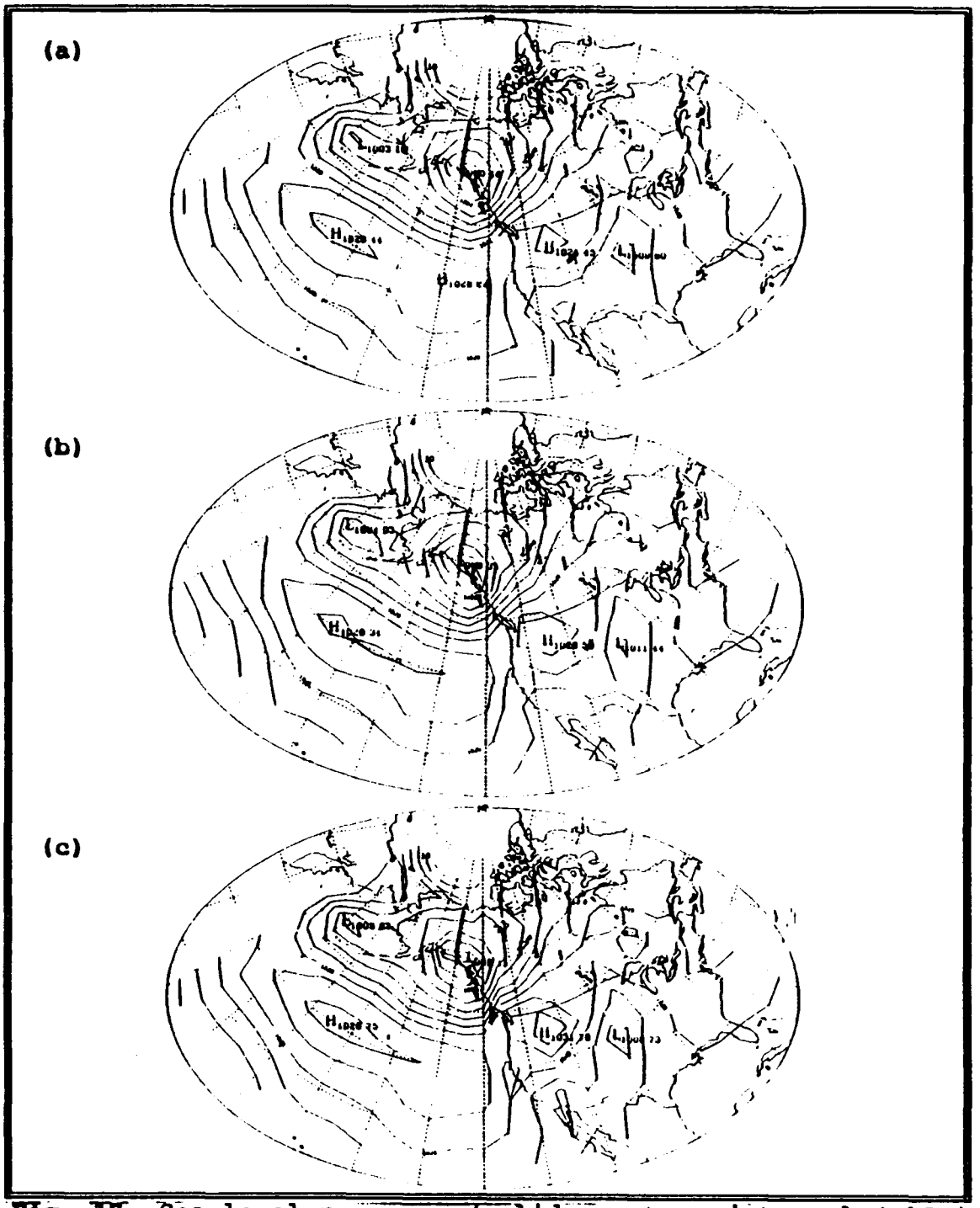

Fig. 17. Sea-level pressure (solid, contour interval 4 hPa) analyses at 1200 UTC 7 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

analysis displayed a higher central pressure value than the NMC or FNOC products, up to approximately 6 hPa higher than the FNOC analysis. In the Gulf of Alaska, central pressure values for the coastal low varied overall by less than 1.5 hPa in the three analyses. Pressure values were also very similar in the eastern Pacific high pressure region, varying by less than 0.5 hPa in the analyses.

2. Surface Pressure Forecast Comparison and Verification

The three major features to be discussed in the forecast comparisons are the Bering Sea low, the Gulf of Alaska coastal low, and the central United States lee cyclone. When verified, the regression forecast provided the best performance in the regions of the Bering Sea low and the lee cyclone. In the Gulf of Alaska, all of the products performed poorly, however, the AVN forecast provided the lowest errors. Details of the forecast comparisons and verification are provided in the following discussion of the synoptic features.

a. Bering Sea Low

In this area, the random nature of the errors in forecast positions and intensities for the individual model forecasts allowed the regression method to provide lower errors in forecasting this feature. The individual model and ensemble surface pressure forecasts for 7 September are shown in Fig. 18a-c and Fig. 19a,b, respectively. Forecast

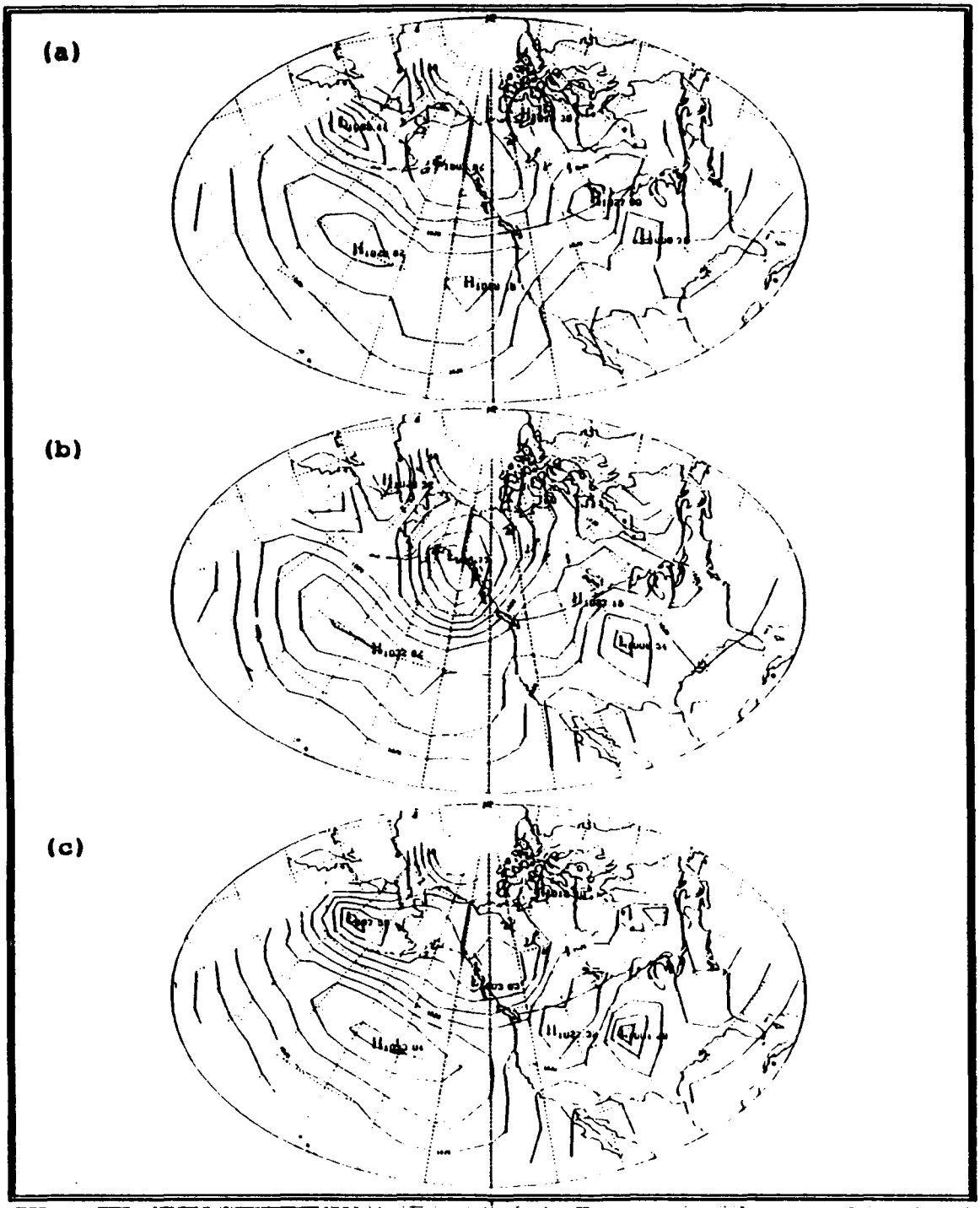

Fig. 18. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 7 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

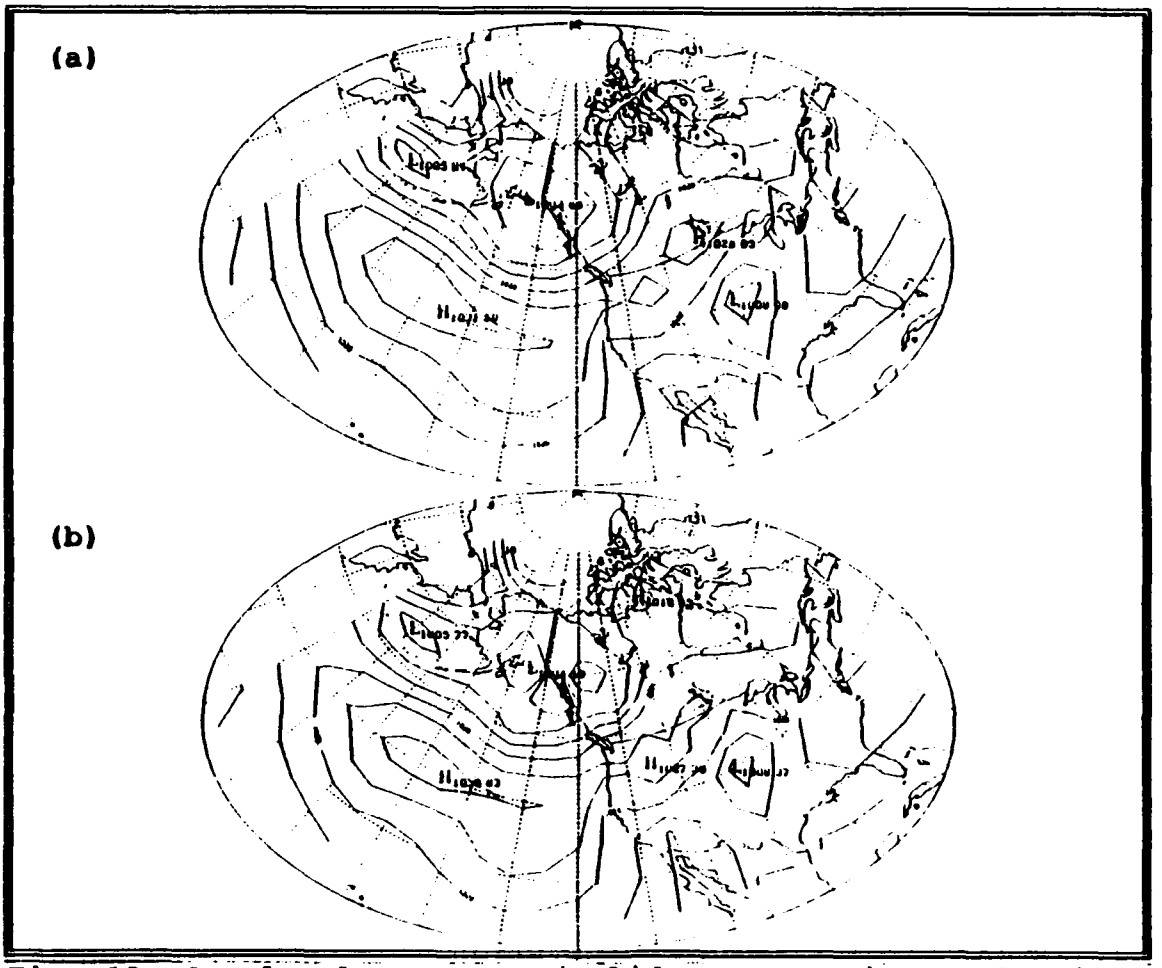

Fig. 19. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 7 September 1992 from the (a) equally weighted and (b) regression ensembles.

differences in both intensity and location were present in the individual models. For example, the FNOC forecast low (Fig. 18a) lagged the ECMWF forecast low (Fig. 18c) in position and carried a central pressure value approximately 5 hPa higher than the ECMWF forecast. The ECMWF low was most accurate in position but was approximately 5 hPa too deep. The FNOC forecast, however, was accurate in intensity but was poor on location. The NMC AVN forecast (Fig. 18b) displayed troughing extending from the west into the area with pressure values roughly 13 hPa too high when compared with the ECMWF analysis. The regressed and equally weighted forecasts (Fig. 19a,b) were accurate in position but were slightly high in pressure when compared to the ECMWF analysis.

The forecast error plots for the individual (Fig. 20a-c) and ensemble (Fig. 21a,b) forecasts showed that the regressed forecast verified the best in the area with errors of 0-3 hPa in the immediate vicinity of the analyzed low. The equally weighted and FNOC forecasts also verified comparably in the area. The AVN forecast displayed a significantly higher error in the area than the remaining products.

b. Gulf of Alaska Coastal Low

In this region, all of the individual forecasts performed poorly. Therefore, the combined influence of the poor individual forecasts degraded the ensemble forecasts.

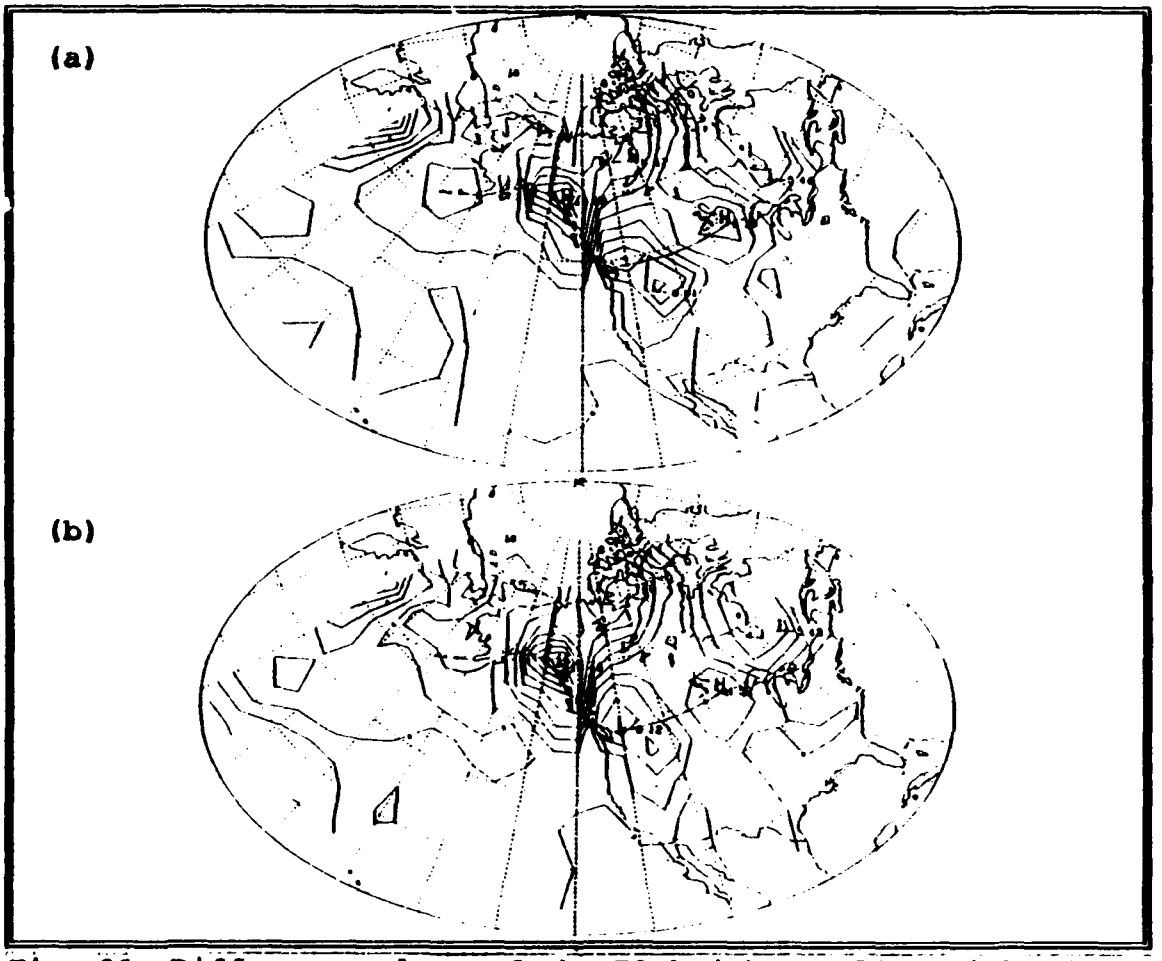

Fig. 21. Difference plots of the 72-h (a) equally weighted and (b) regression sea-level pressure forecasts verified with the ECMWF analysis at 1200 UTC 7 September 1992 (2 hPa contours).

The AVN forecast, although, poor in forecast intensity, provided the lowest overall error.

The forecast position and strength of the low pressure center in this area varied widely in the ECMWF, FNOC, and AVN forecasts. The AVN forecast (Fig. 18b) showed the deepest representation of the low but was more than 7 hPa higher than the analyzed pressure. The ECMWF product (Fig. 18c) showed relatively weak troughing across the area with a relative minimum in pressure forecast well to the south of the analyzed low position in the ECMWF analysis. The regressed and equally weighted forecasts (Fig. 19a,b) placed the low approximately 10° of longitude west of the analyzed position and forecast central pressures approximately 15 hPa higher than the ECMWF analysis.

When verified, all of the forecasts carried pressure values too deep in the area. The AVN forecast showed the lowest error while the remaining forecasts were somewhat comparable to each other. Most of the error in this case consisted of amplitude rather than phase error in the forecasts, although significant phase error was present.

c. Central United States Lee Cyclone

In this region, two of the three individual forecasts provided the proper placement of the low while the forecast pressure values of the individual forecasts were randomly placed about the analyzed value. The combination of

the forecasts in the regression method provided the best overall forecast.

More specifically, the FNOC forecast (Fig. 18a) carried the lee cyclone approximately 5° of longitude east of the low position in the ECMWF analysis. The FNOC and AVN central low pressure values were both approximately 1 hPa higher than the pressure in the ECMWF analysis, while the ECMWF forecast value was approximately 4 hPa lower than the analysis. Differences in the high pressure ridge pattern around the northern side of the low were also present in the individual forecasts. The high pressure values north to northwest of the low were quite similar in the individual model forecasts, however, the FNOC forecast showed a significant weakness in the ridge across the Great Lakes region which was not reflected in the ECMWF analysis. The low pressure and ridge orientation of the equally weighted and regressed forecasts (Fig. 19a,b) closely resembled the ECMWF analysis near the lee cyclone, however, the regressed forecast displayed the greater resemblance to the analysis in the ridge pattern surrounding the low.

The regressed forecast clearly outperformed the other forecasts throughout this region. A broad area of near zero error in the regression forecast verification is shown in Fig. 21b through the central and southeastern United States. The ECMWF and FNOC forecasts overdeveloped the lee cyclone, creating error values of greater than 4 hPa too deep at their

respective forecast low positions. The AVN performed well in the vicinity of the low but was degraded somewhat in surrounding areas. The equally weighted forecast was competitive with the regressed forecast but displayed slightly higher overall error.

3. Comparison of 500-hPa Height Analyses

The major features in the FNOC (Fig. 22a), NMC (Fig. 22b), and ECMWF (Fig. 22c) 500-hPa height analyses are the Bering Sea trough, the northern Gulf of Alaska coastal low, the north central United States trough, and the eastern Pacific high pressure ridge. The FNOC, NMC, and ECMWF 500-hPa height analyses were generally similar in the positioning of synoptic features, however, several differences were noted.

In the eastern Pacific Ocean area, the FNOC analysis displayed the center of the subtropical ridge approximately 10° of longitude farther west of the analyzed position in the ECMWF and NMC analyses. This is not unusual given the data limited nature of the area. Additionally, the NMC analysis showed the subtropical ridge approximately 42 m weaker than the ECMWF and FNOC analyses. In the lower portion of the trough region extending from the north central United States to near the west coast, the NMC and FNOC analyses displayed a sharper trough than the ECMWF analysis. When compared with the trough position in the ECMWF and FNOC analyses, the NMC trough lagged the other analyses in position to the west. In

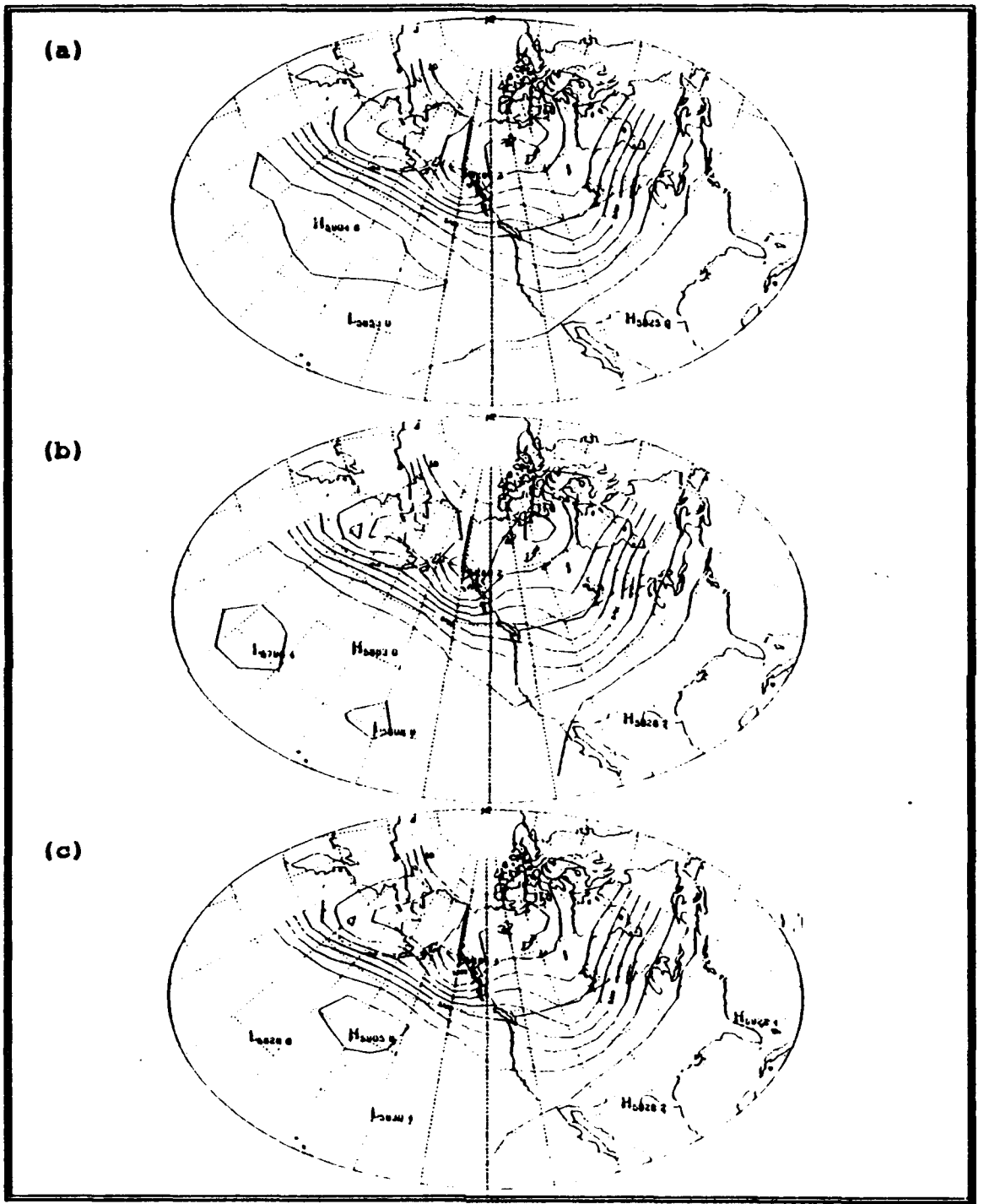

Fig. 22. 500-hPa height (solid, contour interval 60 m) analyses at 1200 UTC 7 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

the Gulf of Alaska coastal area, all three analyses matched on the position of the low. The NMC central height value for the low was, however, approximately 14 m lower than the other analyses. The trough axis orientation in the Bering Sea corresponded well in the analyses. A cutoff low was, however, analyzed by the NMC and ECMWF analyses in the lower portion of the trough. The differences in these analyses would tend to make the ECMWF analysis less desirable or appropriate as a single verification product than in other cases where the analyses corresponded more closely.

4. 500-hPa Height Forecast Comparison and Verification

The performance of the regression forecast in the regions of the major features of this case study is mixed. The ensemble forecasts performed well over the central United States, while selected individual model forecasts performed best in other areas such as the Gulf of Alaska. The following discussion will focus on the forecast comparisons in the regions of the north central United States trough and the Gulf of Alaska/western Canadian low pressure area.

a. North Central United States Trough

In this area, the differences in the individual forecasts of the trough were characterized mainly by somewhat random phase differences. As shown in previous examples of this type of error, the ensemble forecasts generally handle these differences well. In fact, the regressed and equally

weighted forecasts do provide two of the better forecasts in this region.

The individual and ensemble 500-hPa forecasts for 7 September are provided in Fig. 23a-c and Fig. 24a,b, respectively. In general, the ensemble forecasts provided the best placement of the trough when compared to the ECMWF analysis. The FNOC trough was advanced too far eastward while the AVN trough lagged the ECMWF analysis trough position to the west. The ECMWF forecast matched the analysis in the position of the trough axis in the northern portion of the trough but the forecast was too shallow.

The forecast error plots for the 500-hPa forecasts are shown in Fig. 25a-c and Fig. 26a,b. The verified position of the trough fell between the forecast positions of the individual model forecasts. The error patterns of the FNOC and ECMWF forecasts (Fig. 25a,c) were similar to those of the regressed and equally weighted forecasts (Fig. 26a,b), showing a maximum error value to the west of the forecast trough positions in these products. This indicated that the trough in these forecasts was displaced too far to the east. The error pattern of the AVN forecast (Fig. 25b) indicated that the trough had been forecast too far to the west. Overall, the equally weighted forecast appeared to have performed the best in the area, however, the regressed and FNOC forecasts were comparable. The ECMWF product showed the highest error value associated with trough displacement and intensity error.

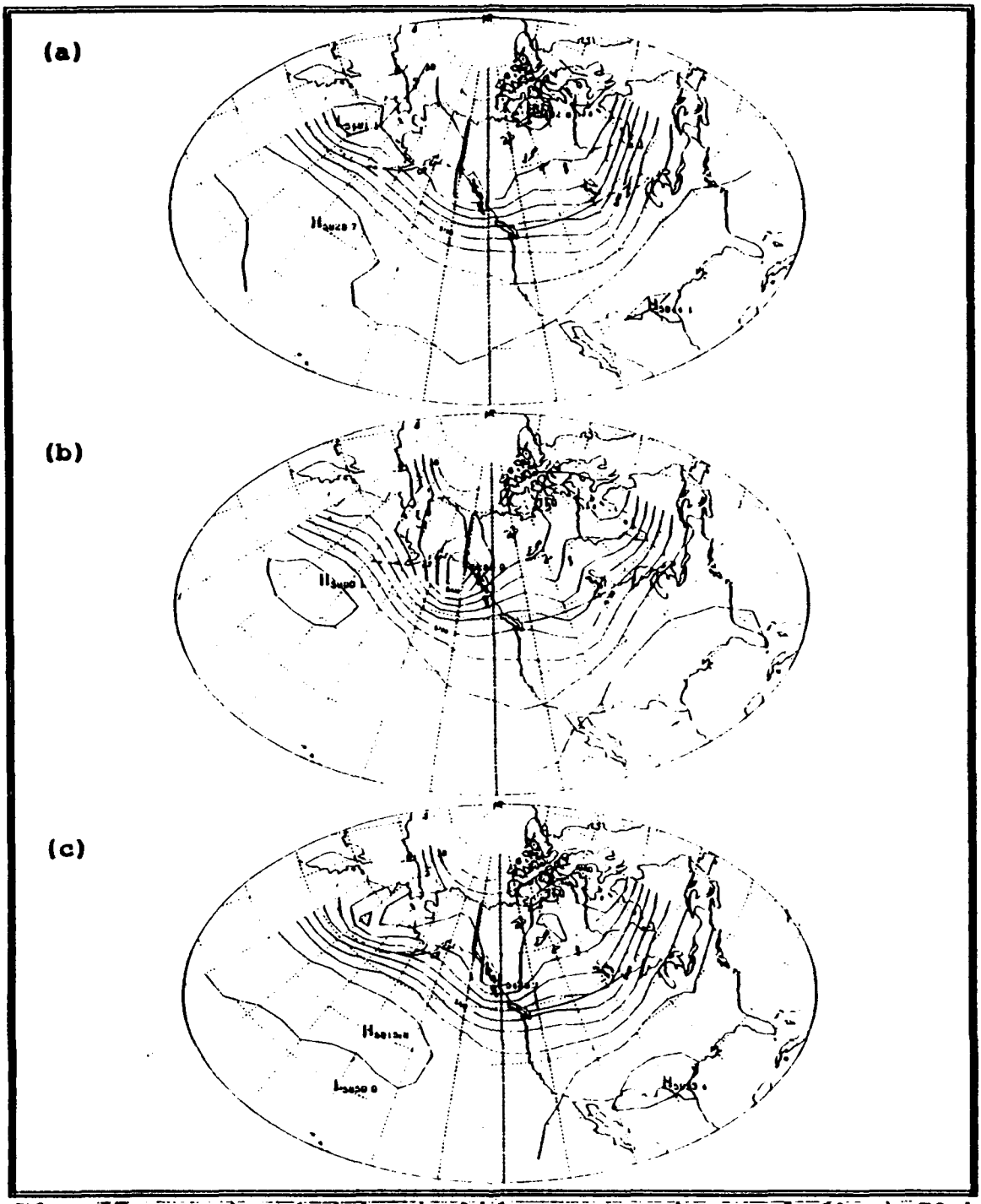

Fig. 23. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 7 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

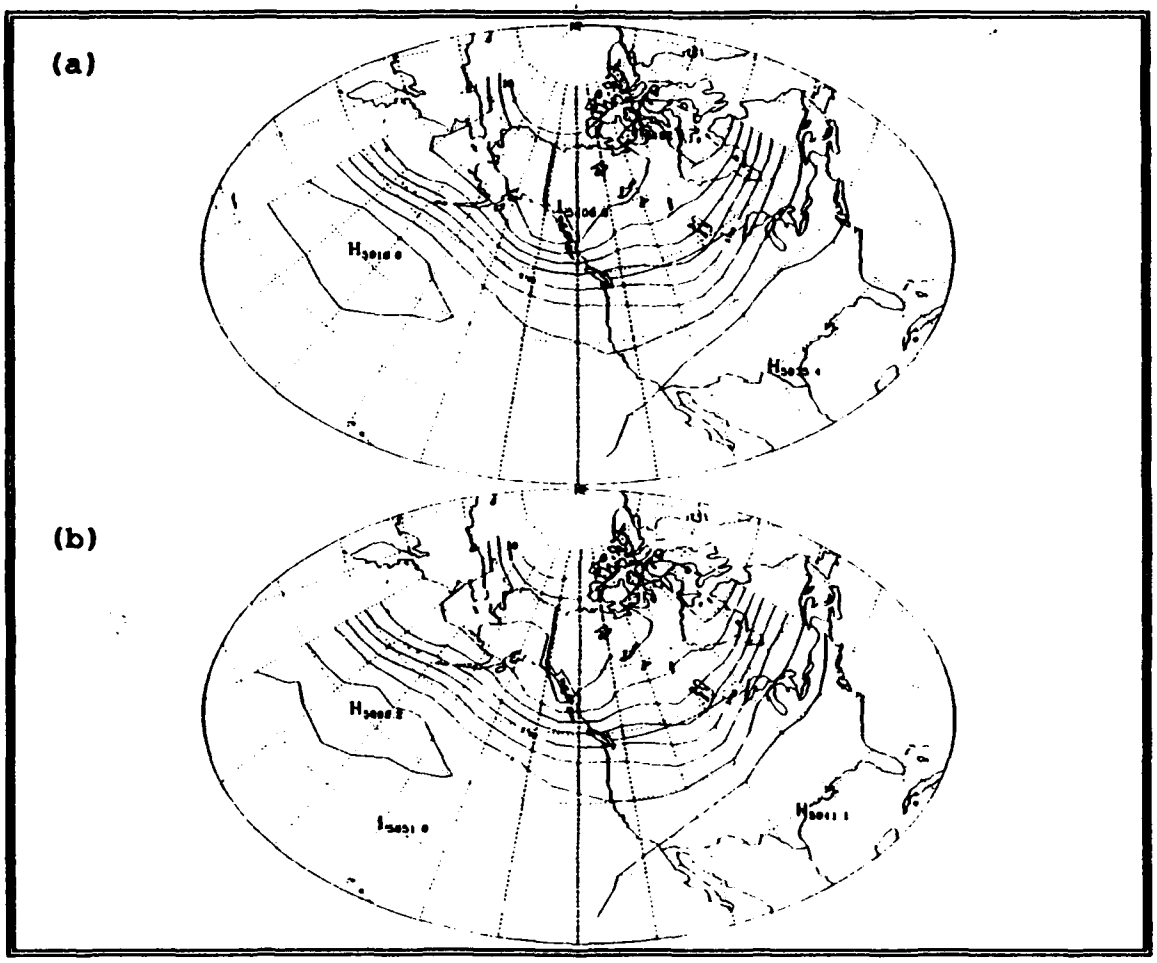

Fig. 24. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 7 September 1992 from the (a) equally weighted and (b) regression ensembles.

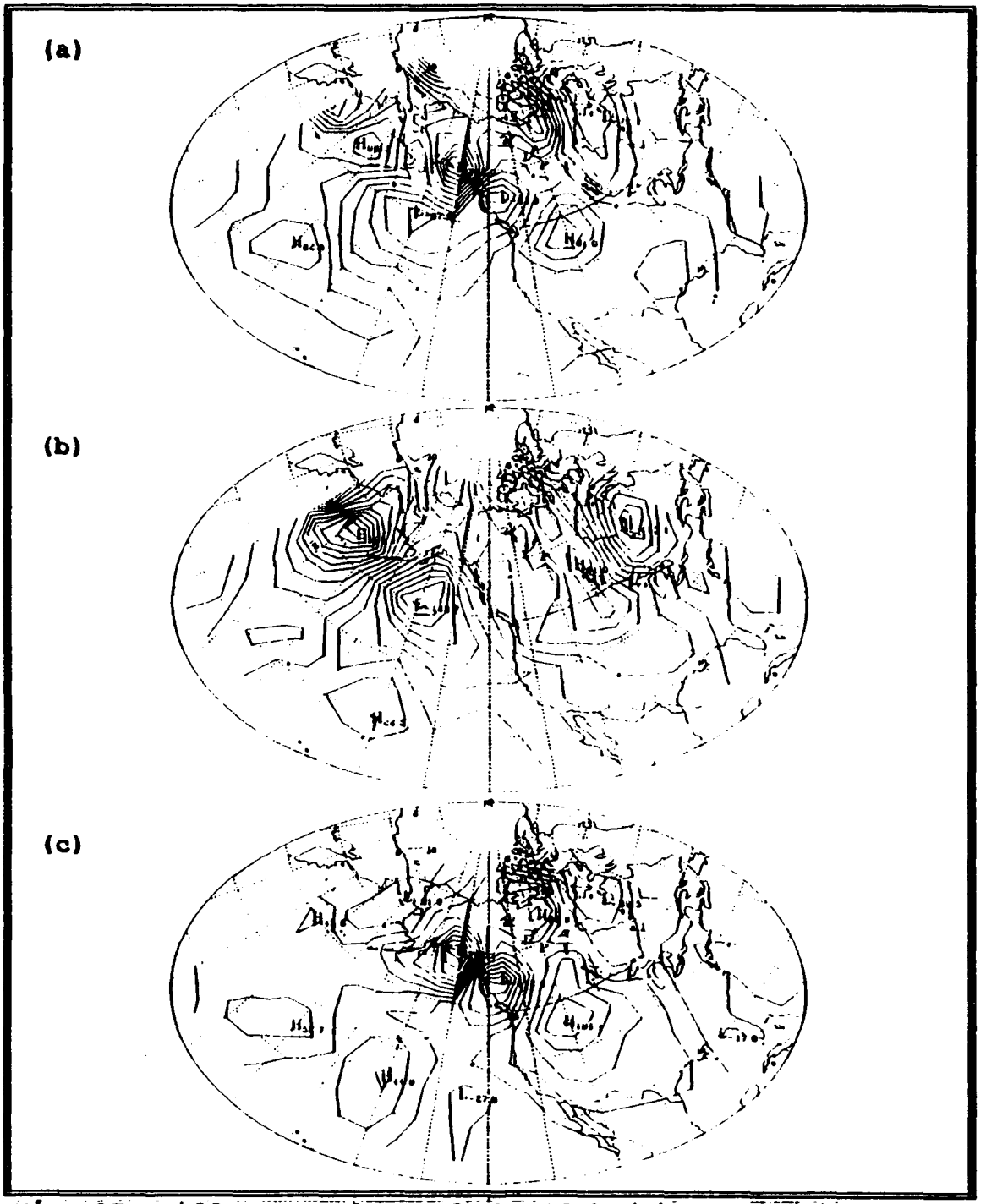

Fig. 25. Difference plots of the 72-h (a) FNOC, (b) NMC AVN, and (c) ECMWF 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 7 September 1992 (20 m contours).

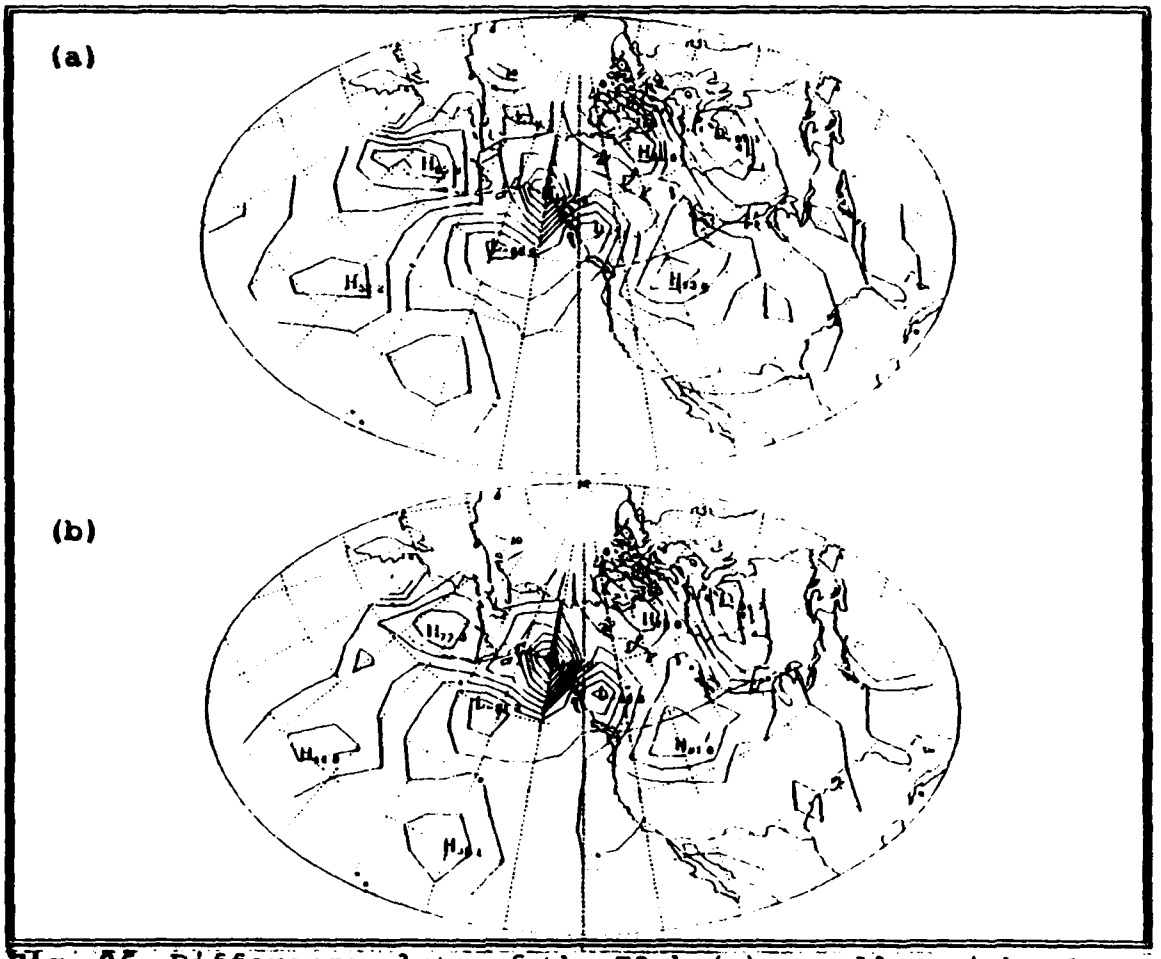

Fig. 26. Difference plots of the 72-h (a) equally weighted and (b) regression 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 7 September 1992 (20 m contours).

b. Gulf of Alaska/Western Canadian Low Pressure Area

In this region, only the AVN forecast captured the low pressure area which verified in the area. The remaining forecasts missed the low and performed poorly. The poor performance by these forecasts greatly degraded the performance of the ensemble forecasts.

The AVN product (Fig. 23b) accurately forecast a significantly deeper low and associated southward extending trough than the other forecast products in the area. The central height value of the low on the AVN product was approximately 127 m deeper than the ECMWF low and greater than 100 m deeper than height values on the FNOC forecast, however, the AVN height value was within 2 m of the central height value on the ECMWF analysis. Additionally, the center of the AVN low coincided directly in position with the ECMWF analysis low position. The deeper 500-hPa AVN trough corresponded with the deeper surface low forecast in the area by the AVN surface forecast.

Associated with the deeper AVN low, a strong height gradient was indicated on the south side of the low which was not as intense in the other products but did match well with the ECMWF analysis. Additionally, a sharper ridge to the east of the trough was indicated on the AVN forecast which also coincided with the ECMWF analysis.

The equally weighted and regressed forecast products (Fig. 24a,b) inaccurately exhibited a shallower low and a weaker gradient to the south of the low than the ECMWF analysis. These differences represented the moderating effects of the ECMWF and FNOC forecast inputs which degraded the performance of the ensemble forecasts in this case.

The error plots in Fig. 25a-c and Fig. 26a,b indicated that the AVN forecast verified quite well in the vicinity of the low, while the ECMWF forecast demonstrated the weakest performance in the area. The regressed forecast demonstrated the best performance in the trough area to the southwest of the low. The AVN forecast verified best in the ridge to the east of the low, while the regressed forecast outperformed the remaining products in this area.

D. CASE 3 (1200 UTC 15 SEPTEMBER, 1992)

Case 3 is similar to Case 2 in terms of the presence of a lee cyclone and is similar to both previous cases in terms of the presence of a Gulf of Alaska low pressure feature. However, this case differs from the previous cases in the eastern Pacific Ocean where the subtropical high has become somewhat displaced by the presence of two mid-latitude low pressure areas.

1. Comparison of Surface Pressure Analyses

The major features in the FNOC (Fig. 27a), NMC (Fig. 27b), and ECMWF (Fig. 27c) sea-level pressure analyses are the

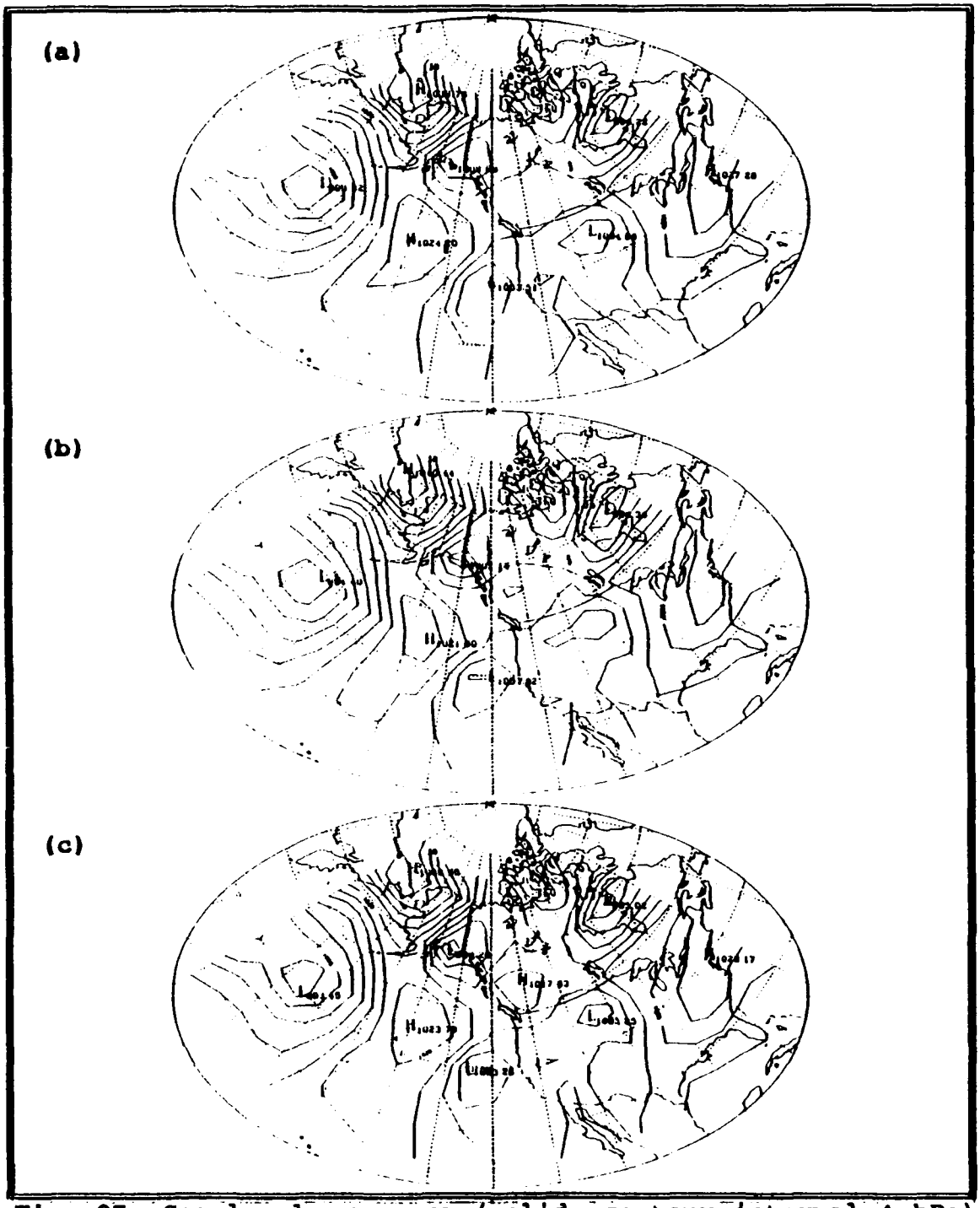

Fig. 27. Sea-level pressure (solid, contour interval 4 hPa) analyses at 1200 UTC 15 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

Gulf of Alaska coastal low, the central United States lee cyclone, the eastern Pacific low near the west coast of California, the subtropical high in the eastern Pacific, and the low pressure area south of the Aleutian Islands. Comparison of the analyses for these features showed that the NMC low pressure area in the northern Gulf of Alaska was analyzed approximately 5° of longitude east of the corresponding low pressure centers indicated on the ECMWF and FNOC analyses. The ECMWF analysis exhibited a central pressure for the low approximately 4 hPa deeper than the FNOC low and 2 hPa deeper than the NMC low. Analyzed positions of the lee cyclone in the central United States matched closely in the three analyses. The central pressure of the feature in the analyses also matched to within 0.5 hPa. The positions of the high and low pressure areas in the eastern Pacific Ocean matched well in the three analyses. The central pressures of both low pressure areas in the three analyses also matched to within 1 hPa. A slightly higher difference in analyzed pressures was noted for the subtropical high in the area. The NMC product analyzed the central pressure within the subtropical high in the eastern Pacific Ocean at approximately 2-2.5 hPa higher than the ECMWF and FNOC analyses. Overall, these analyses coincided well and the use of the ECMWF analysis as the verification product is appropriate.

2. Surface Pressure Forecast Comparison and Verification

Comparison of the individual (Fig. 28a-c) and ensemble (Fig. 29a,b) forecasts through difference plots (Fig. 30a-c and Fig. 31a,b) indicated that the regression forecast was the best product in forecasting the features of this case study at the surface. The following discussion will provide details on the forecast comparisons of the Gulf of Alaska/western Canadian low pressure area, the central United States lee cyclone, and the eastern Pacific Ocean synoptic features.

a. Gulf of Alaska/Western Canadian Low Pressure Area

In this area, the individual model forecast low positions were scattered around the analyzed position of the low and were all too deep in their forecast central pressures. However, the blending of these forecasts in the regression method provided the best forecast of the low which was very accurate in intensity and similar in location.

For comparison, the individual and ensemble surface pressure forecasts for 15 September are provided in Fig. 28 a-c and Fig. 29a,b, respectively. The intensity, orientation, and position of the low pressure in the western Canada/Gulf of Alaska area were forecast significantly differently in the FNOC, AVN, and ECMWF forecasts (Fig. 28a-c). Of these forecasts, the FNOC product forecast the highest central pressure at 990.40 hPa while the AVN product forecast the lowest pressure at 986.70 hPa. The FNOC and AVN forecasts

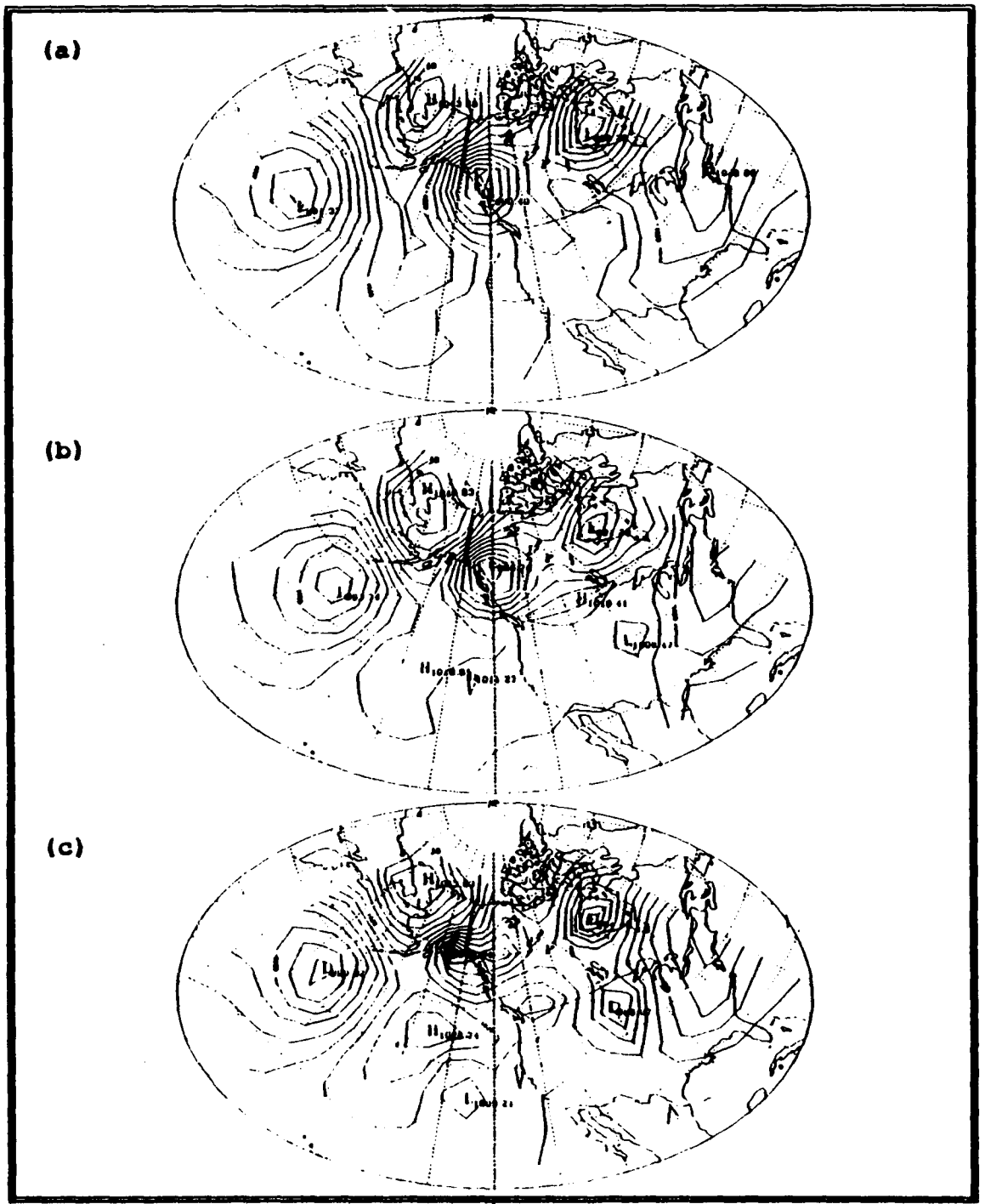

Fig. 28. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 15 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

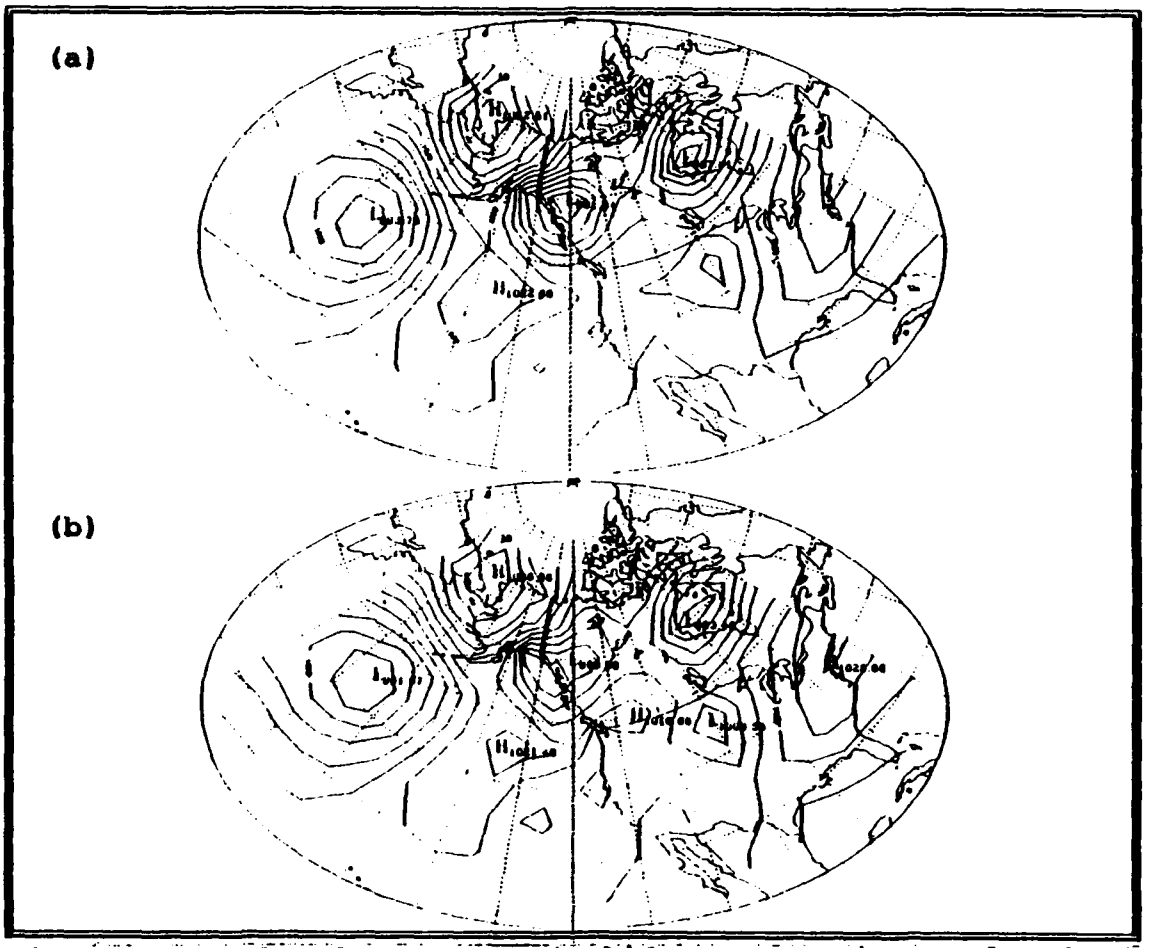

Fig. 29. Sea-level pressure (solid, contour interval 4 hPa) 72-h forecasts at valid time 1200 UTC 15 September 1992 from the (a) equally weighted and (b) regression ensembles.

displayed a single low pattern while the ECMWF showed the low in a split pattern not indicated on the ECMWF analysis, possibly due to topographic effects on the forecast for the area. The regressed forecast (Fig. 29b) carried a broad area of low pressure across the area with a significantly higher and more accurate forecast central pressure than the individual forecasts. The equally weighted forecast (Fig. 29a) indicated a central pressure much closer to the values in the individual forecasts. The orientation of the low in the equally weighted forecast was similar to the pattern of the AVN and FNOC lows reflecting the effects of the equal contributions of the three individual forecasts.

As shown in the error plots (Fig. 30a-c and Fig. 31a,b), the regressed forecast displayed the best performance in the area. The ECMWF forecast error values (Fig. 30c) were also comparable. The FNOC, AVN, and equally weighted forecast errors were significantly higher (Fig. 30a,b and Fig. 31a). The large magnitude high/low pattern of the error in the AVN and FNOC error plots reflected the phase error of the forecast low pressure center in these products.

b. Central United States Lee Cyclone

In the lee cyclone area, the regression forecast provided the best representation of the feature. The representations of this feature by the individual forecasts

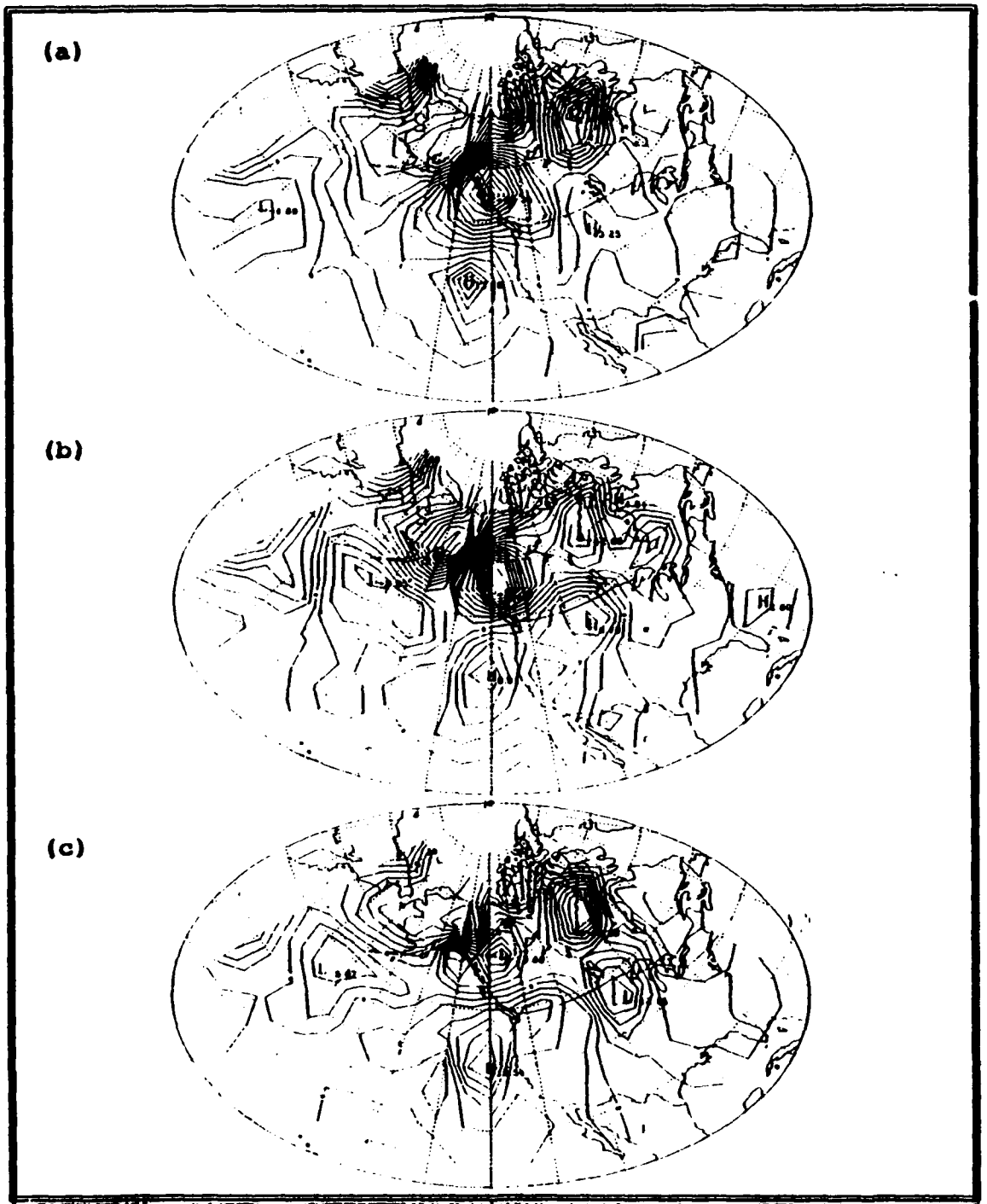

Fig. 30. Difference plots of the 72-h (a) FNOC, (b) NMC AVN, and (c) ECMWF sea-level pressure forecasts verified with the ECMWF analysis at 1200 UTC 15 September 1992 (2 hPa contours).

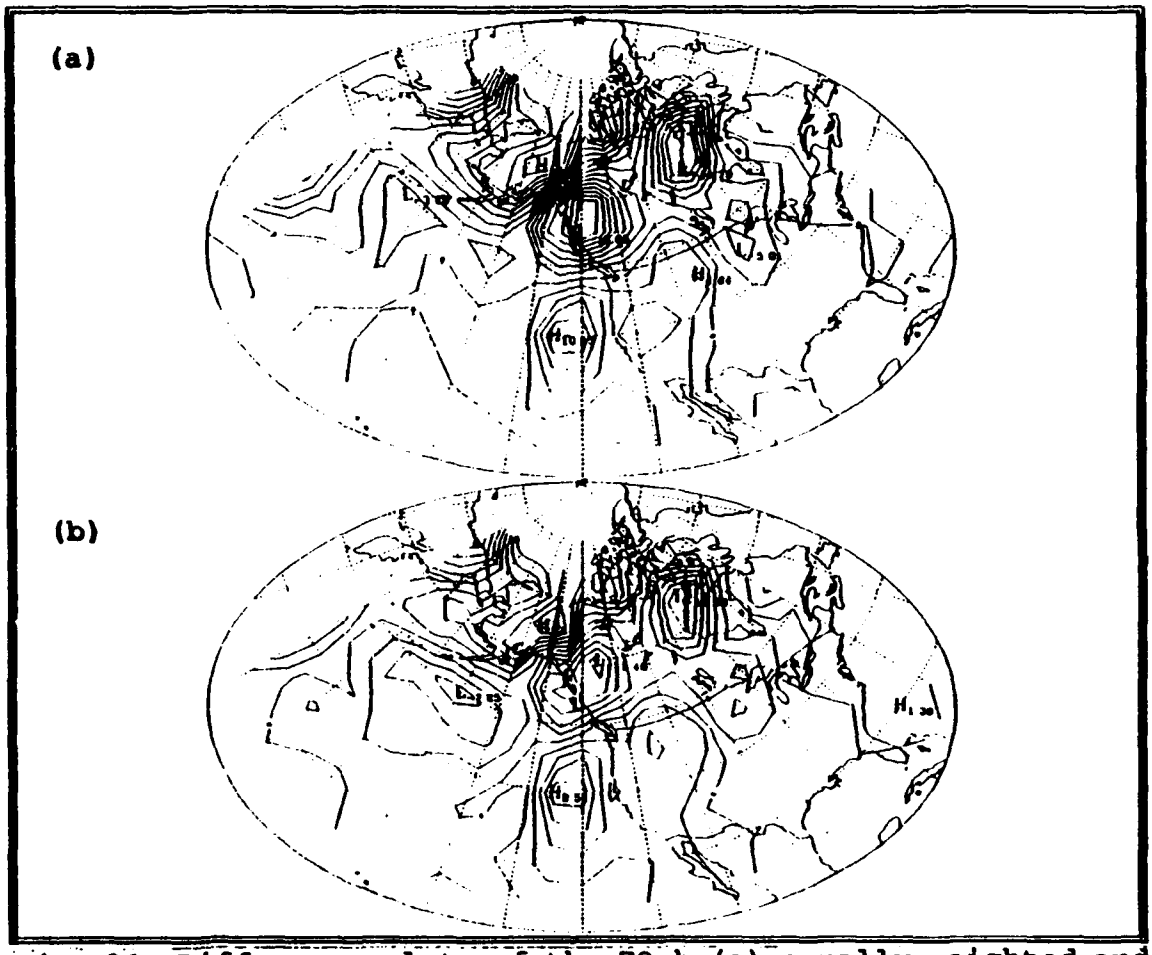

Fig. 31. Difference plots of the 72-h (a) equally weighted and (b) regression sea-level pressure forecasts verified with the ECMWF analysis at 1200 UTC 15 September 1992 (2 hPa contours).

varied greatly and the errors by each were random in nature. This contributed to the success of the regression forecast.

Comparison of the individual forecasts revealed that the ECMWF forecast (Fig. 28c) showed the deepest representation of a developing lee cyclone in the area approximately 7 hPa deeper than that of the ECMWF analysis. The AVN product (Fig. 28b) forecast the low approximately 3 hPa shallower and approximately 5° of latitude farther southeast than the ECMWF analysis. The FNOC product (Fig. 28a) forecast only an area of troughing across the area with pressure values approximately 3-7 hPa shallower than the ECMWF analysis in the area. The regressed forecast (Fig. 29b) placed the lee cyclone slightly east of the analyzed position and was within 2 hPa of the analyzed central pressure. The equally weighted forecast (Fig. 29a) matched the regressed forecast in position and was similar in pressure.

The error plot of the regressed forecast (Fig. 31b) reflects the low overall error of the regressed product in the area. The error plots of the FNOC and equally weighted forecasts (Fig. 30a and Fig. 31a) indicated that these products were also comparable in the region.

c. Eastern Pacific Ocean Synoptic Features

A general heading has been applied to this section in order to provide a description of the forecast differences for a set of interconnected synoptic features extending across

the area. These features include a low pressure area near the dateline, a subtropical high pressure region to the east, and a second low pressure area directly west of the central California coast. Within this region, the regression forecast provided the best overall performance across the region on average. The successful performance of the regression forecast was assisted by the random nature of the errors in the individual model forecasts of these features.

For example, the individual model forecasts of the western low pressure area differed somewhat randomly both in position and strength. A maximum difference of approximately 5° of latitude in forecast positions existed between the individual model forecast positions of the low and the analyzed position, while a maximum difference of approximately 3.5 hPa in forecast central pressures existed between the ECMWF forecast and the analyzed pressure. The FNOC forecast matched the analysis in the position of the low while the AVN forecast was very similar to the analysis in terms of central pressure for the low. The equally weighted and regressed forecasts (Fig. 29a,b) positioned the low at an intermediate position which reflected the blending of the individual model forecast low positions and was 4°-5° of latitude north of the analyzed position. The regressed and equally weighted products forecast intermediate central pressure values which were within 2 hPa of the analyzed pressure.

The individual model forecasts displayed significant differences in the subtropical high and the associated ridge extending to the northeast. Central pressure values were approximately 3 hPa too high in the ECMWF forecast and approximately 4 hPa too low in the AVN forecast. The FNOC forecast inaccurately displayed a significantly weaker ridge pattern to the northeast than was shown in the other forecasts. The regressed and equally weighted products forecast central pressures in the high to within 2 hPa of the analyzed value.

The ECMWF and AVN forecasts indicated the low pressure area west of California which was represented by only weak troughing in the FNOC product. However, the ECMWF and AVN forecasts were 6-8 hPa too shallow in the forecast of the central pressure in the low while the pressure values of the FNOC forecast were approximately 13 hPa too high. The regressed and equally weighted forecasts also carried weak areas of low pressure in this area which were approximately 7 hPa too high in magnitude and were displaced approximately 5°-10° of longitude east of the analyzed position. This represented a case where all of the individual forecasts were poor, therefore the regressed forecast was somewhat degraded but still displayed improvement in general over the individual forecasts.

When verified, the regressed forecasts provided the lowest errors in the eastern portion of the region (Fig. 30b

and Fig. 31b). Although error comparison is more difficult in the western area without focusing on specific features, the regressed forecast appeared to provide the lowest error on average in the area. Several of the remaining forecasts were comparable with the regressed forecast in the western region.

3. Comparison of 500-hPa Height Analyses

The major features in the FNOC (Fig. 32a), NMC (Fig. 32b), and ECMWF (Fig.32c) 500-hPa height analyses are the eastern Pacific Ocean low, the eastern Pacific subtropical ridge, the high-amplitude trough near the west coast of the United States, and the low over eastern Alaska and western Canada.

The analyses were close in comparison overall, however, differences were noted. In the analysis of the 500-hPa low over eastern Alaska and western Canada, the NMC analysis showed the low to be significantly northeast of the positions in the FNOC and ECMWF analyses. Additionally, the NMC product exhibited the lowest central height of the three analyses, approximately 26 m deeper than the ECMWF analysis. In the associated trough pattern extending southward from the low, the NMC analysis was approximately 20-30 m deeper than the other analyses in the trough across Canada. In the southern region of the trough west of California, the NMC and ECMWF analyses carried the trough up to approximately 40 m deeper than the FNOC analysis. The three analyses matched

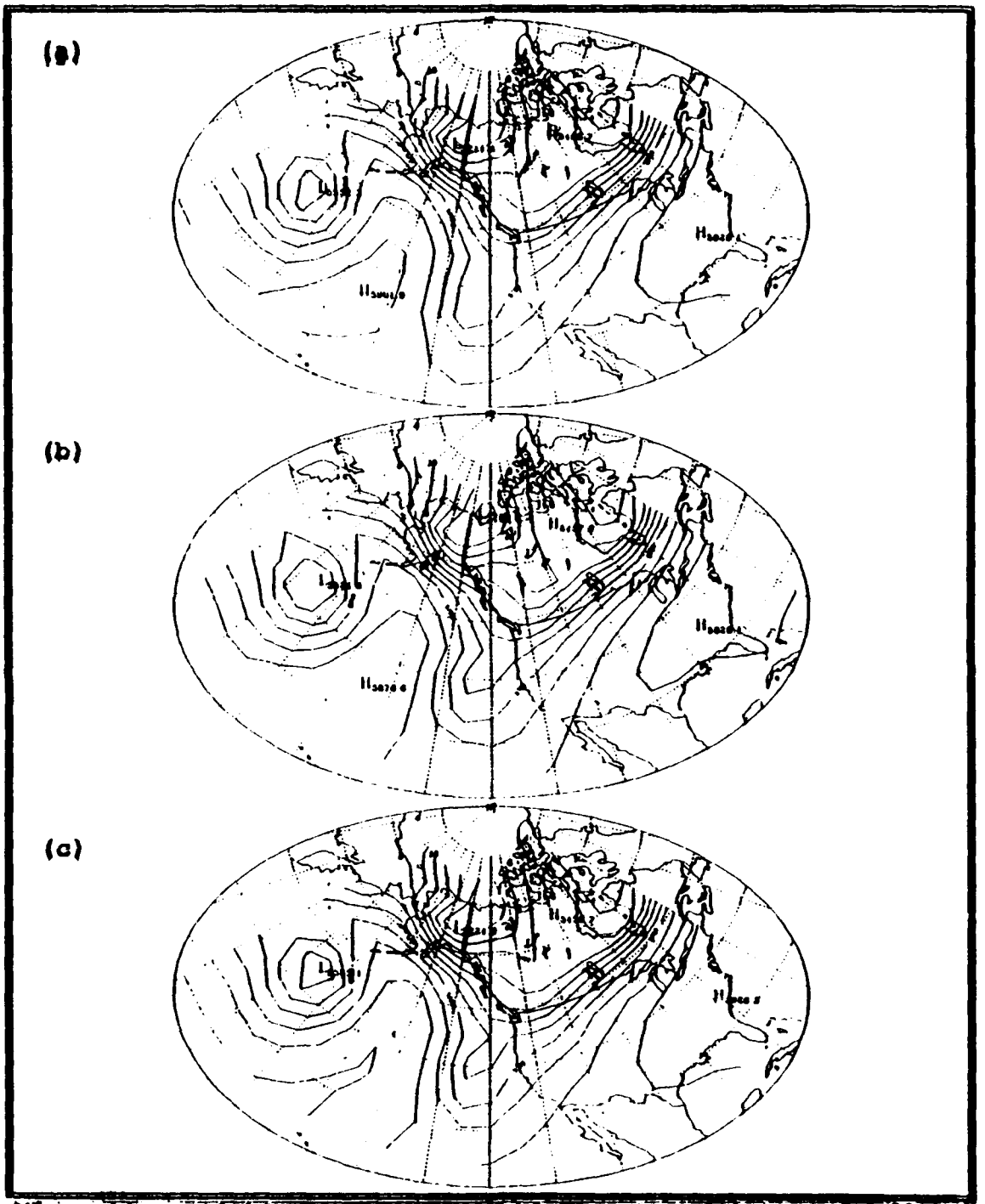

Fig. 32. 500-hPa height (solid, contour interval 60 m) analyses at 1200 UTC 15 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

well in the orientation and positions of the eastern Pacific low and subtropical ridge, however, height differences were present. The analyses varied by as much as 30 m in the low and by 25 m in the ridge. The ECMWF analysis would be less desirable as a single verification product in this case given the significant differences present in several of the analyzed features.

4. 500-hPa Height Forecast Comparison and Verification

Comparison of individual (Fig. 33a-c) and ensemble (Fig. 34a,b) forecasts through difference plots (Fig. 35a-c and Fig. 36a,b) revealed mixed results of regression forecast performance. In general, the forecasts of the major features in this case are poor, however, the regressed forecast does show gains over the individual forecasts in forecasting the low over eastern Alaska and western Canada. Details of the forecast comparisons in this area as well as in the trough feature along the coast of the western United States and the weak trough feature in the northern United States are provided in the following discussion.

a. Western Canadian Low

The individual model forecasts displayed consistently poor performances in this area. Therefore, the ensemble forecasts were degraded due to these weak forecast contributions. However, the regression forecast did provide the greatest reduction in forecast error in the area.

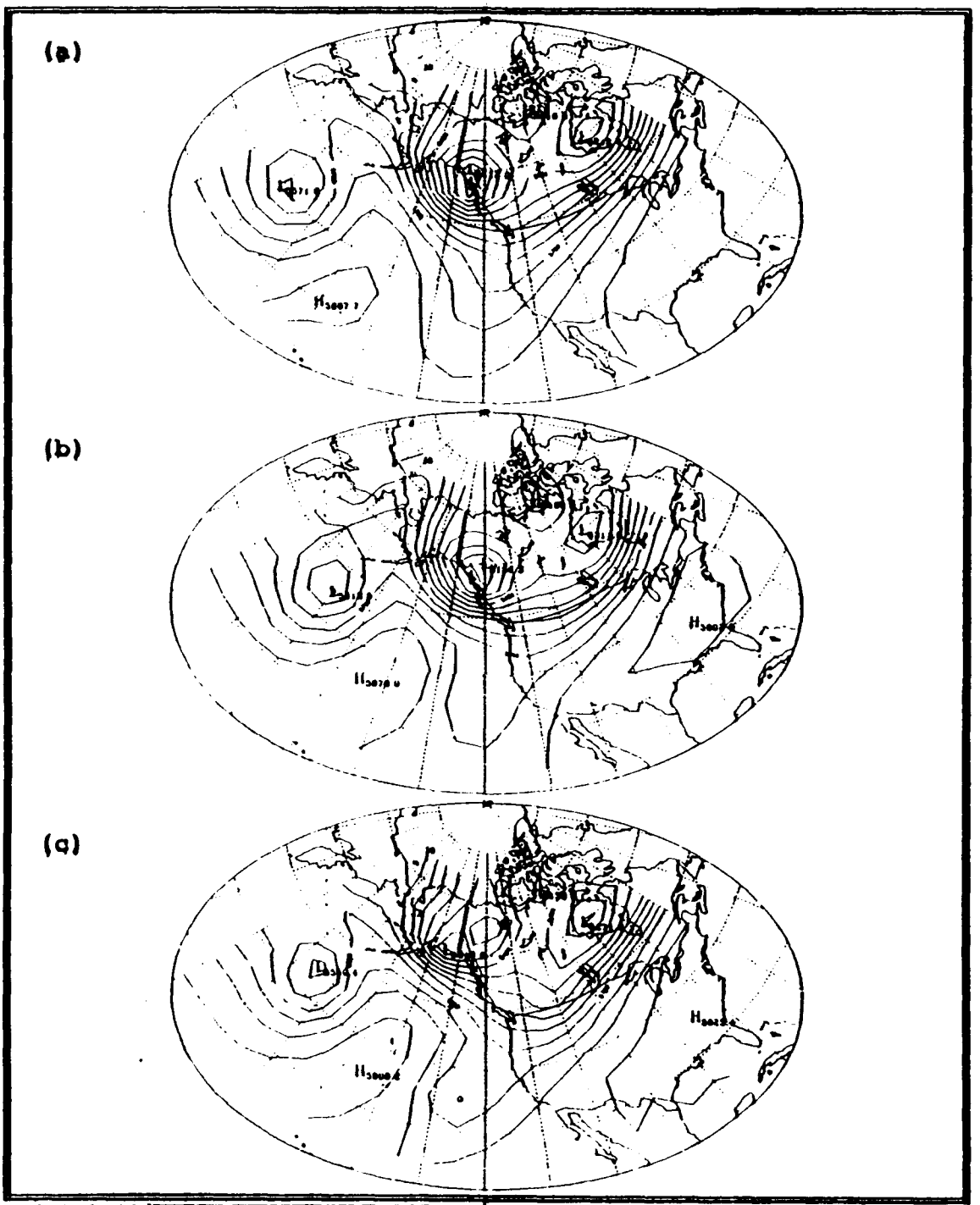

Fig. 33. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 15 September 1992 from the (a) FNOC, (b) NMC, and (c) ECMWF models.

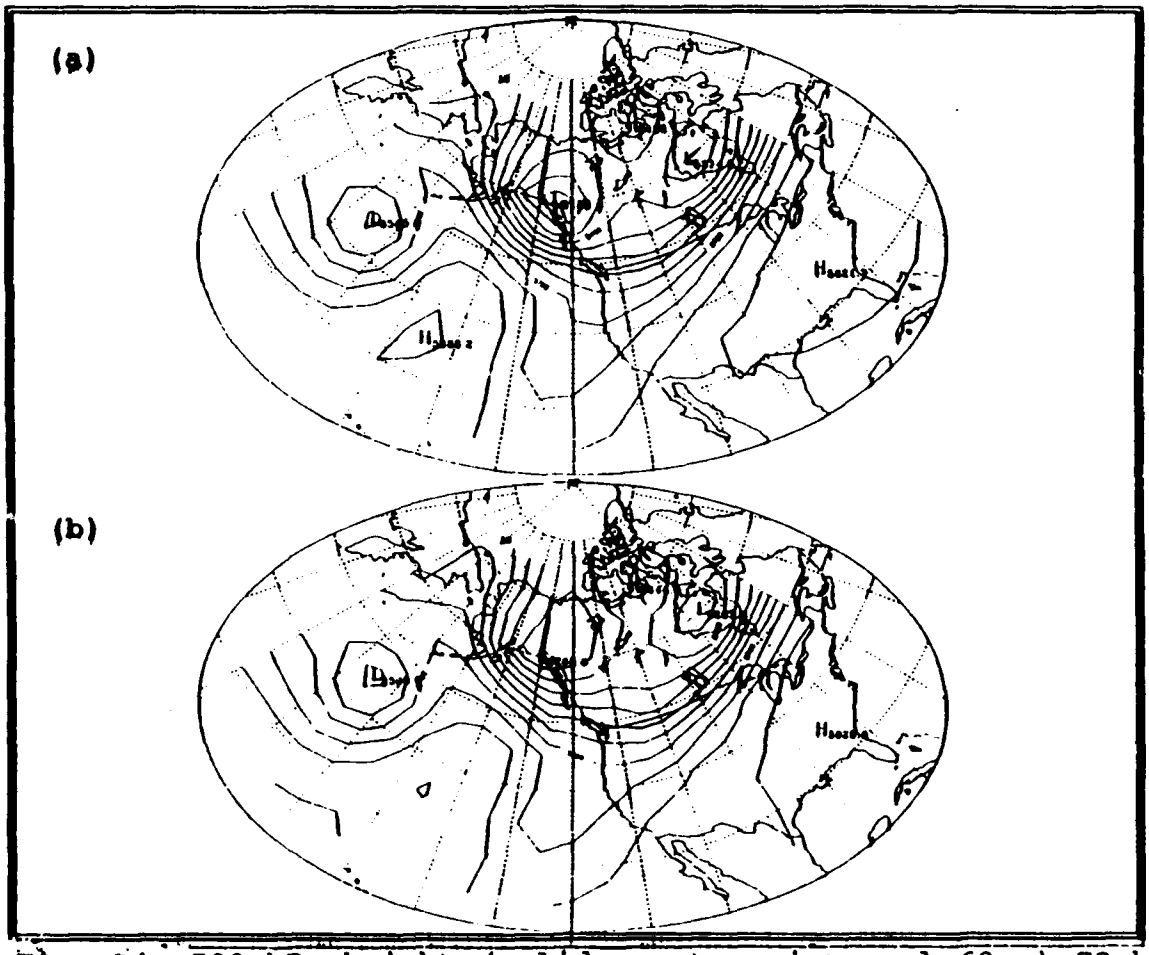

Fig. 34. 500-hPa height (solid, contour interval 60 m) 72-h forecasts at valid time 1200 UTC 15 September 1992 from the (a) equally weighted and (b) regression ensembles.

As an example of the error in the individual model forecasts of this feature, the low positions in the individual forecasts were forecast up to 15° of longitude away from the analyzed position and were generally south to southeast of this position. The AVN product (Fig. 33b) forecast an eastern position in western Canada, while the ECMWF product (Fig. 33c) forecast a dual low pattern with a western center forecast over eastern Alaska. The FNOC forecast position was at an intermediate position (Fig. 33a). The central height of the low was forecast up to 110 m too deep by the individual forecasts. Both the regressed and equally weighted products (Fig. 34a,b) forecast intermediate positions for the low, with the regressed forecast exhibiting the most accurate position. The regressed forecast produced the highest central height value for the low of any of the forecasts.

Although none of the forecasts verified extremely well in the area as seen in the error plots in Fig. 35a-c and Fig. 36a,b, the regressed forecast provided the lowest overall error. The dual high/low pattern of the error plots indicated the low was generally forecast too far to the southeast of the verified location.

b. Western United States Coastal Trough

All of the individual forecasts performed poorly in this area with large amplitude errors. The ensemble forecasts

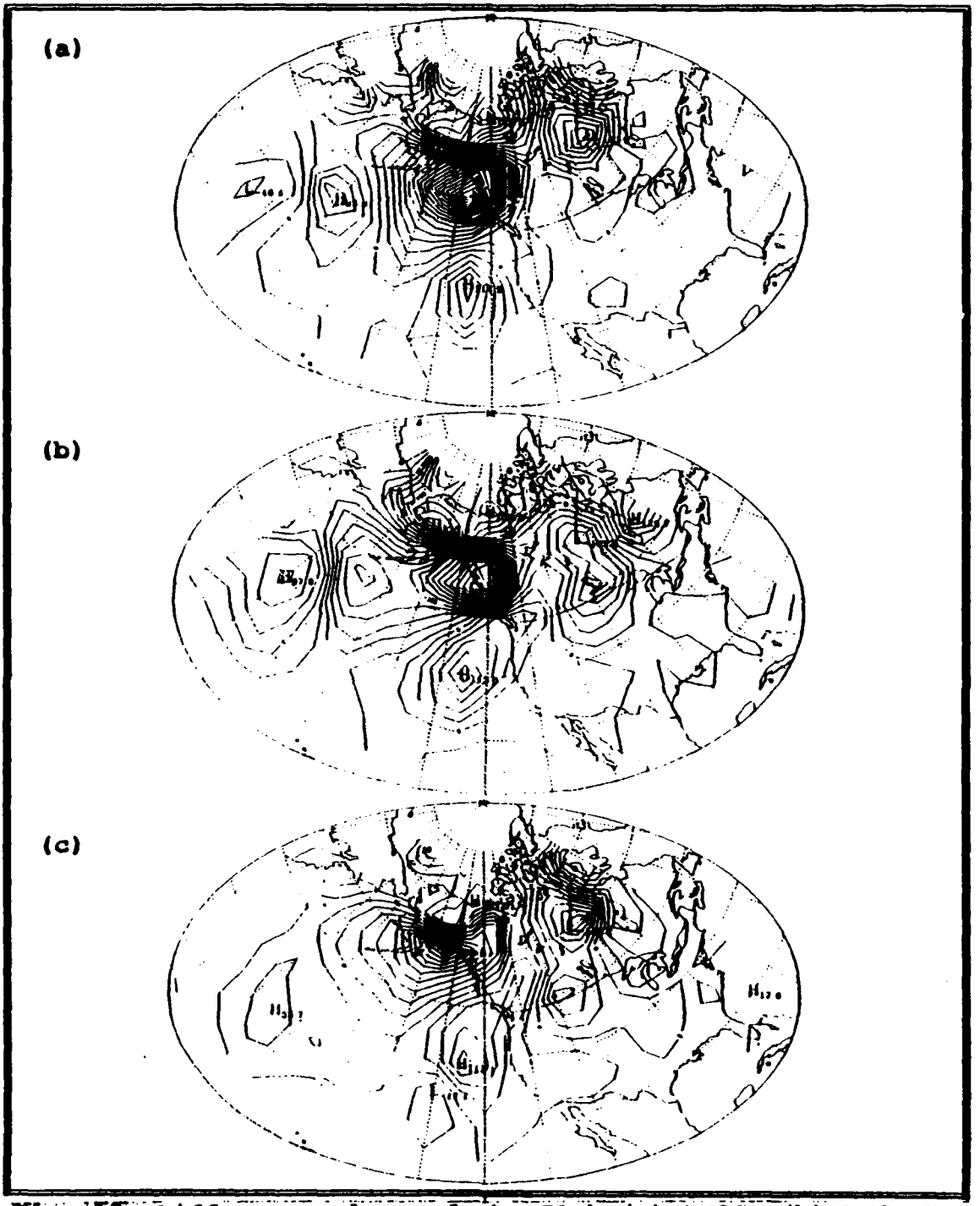

Fig. 35. Difference plots of the 72-h (a) FNOC, (b) NMC AVN, and (c) ECMWF 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 15 September 1992 (20 m contours).

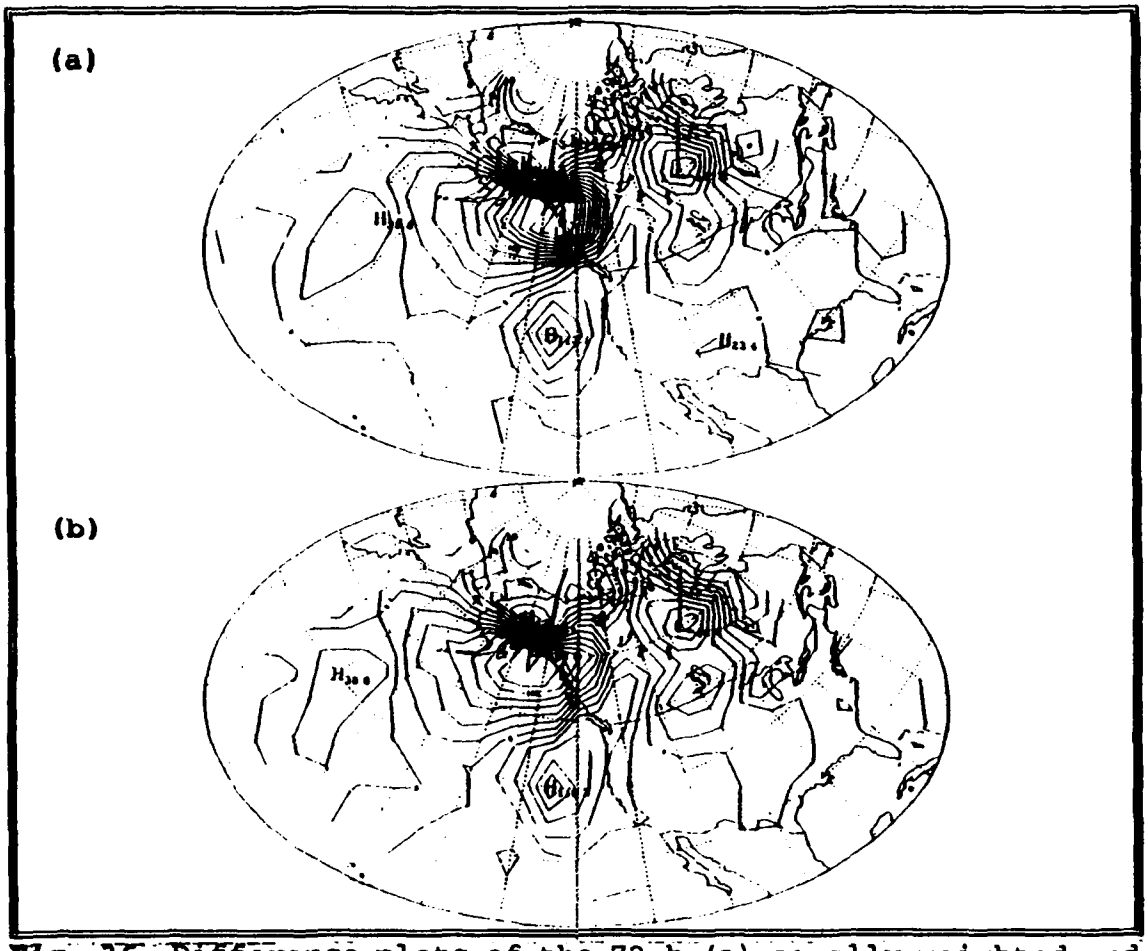

Fig. 36. Difference plots of the 72-h (a) equally weighted and (b) regression 500-hPa height forecasts verified with the ECMWF analysis at 1200 UTC 15 September 1992 (20 m contours).

were therefore also degraded. All of the forecasts produced large errors in the area which were comparable in magnitude.

Comparison of the forecasts revealed that the ECMWF and AVN forecasts of this feature were similar, however, the southern end of the trough in the ECMWF forecast (Fig. 33c) was located west of the AVN position and matched most accurately with the analysis. As in the AVN forecast, the FNOC trough position was also east of the analyzed trough position. The FNOC forecast (Fig. 33a) displayed the deepest representation of the trough but was still approximately 70 m shallower than the analyzed height in the trough. The regressed and equally weighted forecasts (Fig. 34a,b) both displayed trough axes slightly east of the analyzed axis, however, both forecasts were 60-80 m shallower than the analysis in intensity.

When verified, the upper trough was forecast too deep in the FNOC product (Fig. 35a) and too shallow in the remaining products. The average error for each of the forecasts in the region of the trough was quite similar, indicating that the trough, in general, had not been forecast deep enough by any of the forecasts.

c. North Central United States Trough

Although this feature did not appear in every forecast, its association with the developing lee cyclone at the surface added significance to its investigation. The FNOC

forecast performed the best in the area by exhibiting no indication of a trough in the region. A weak erroneous indication of the trough in the AVN and ECMWF forecasts slightly degraded the ensemble forecasts.

The ECMWF forecast (Fig. 33c) showed the sharpest representation of a trough in the area while the AVN forecast (Fig. 33b) displayed only a slight indication of the trough. More accurately, the FNOC forecast (Fig. 33a) did not show any indication of a trough in the area. The regressed and equally weighted forecasts (Fig. 34a,b) did show an indication of the trough in the area as a result of the AVN and ECMWF forecast contributions, although not as sharply defined as the ECMWF forecast.

The FNOC product verified very well in the area as shown in the error plots (Fig. 35a-c and Fig. 36a,b). The equally weighted and regressed forecasts, although not as accurate as the FNOC forecast, outperformed the ECMWF and AVN forecasts. The verification indicated that, in fact, there was minimal 500-hPa support at this time period for the developing surface lee cyclone.

d. Eastern Pacific Ocean Low

As in many of the previous cases, the random nature of the errors in the individual forecasts of this feature contributed to successful ensemble forecasts. As a result,

the regressed forecast provided the best performance in this area.

Differences in position and intensity of the low were noted in the FNOC, AVN, and ECMWF forecasts (Fig. 33a-c). A maximum phase difference of approximately 10° of longitude was noted between the AVN and FNOC forecasts, while a maximum amplitude difference of approximately 59 m was noted between the same two forecasts. The ECMWF position matched the analyzed position of the low while the FNOC and AVN forecast positions were well southwest and east of the analyzed position, respectively. The regressed and equally weighted forecasts (Fig. 34a,b) accurately matched the analyzed low pressure center, but forecast central heights 40-45 m shallower than the analyzed heights.

The error plots for the forecasts (Fig. 35a-c and Fig. 36a,b) indicated that the regressed forecast had the lowest overall error in this region. The equally weighted forecast, however, was comparable and both ensemble forecasts showed significant improvement over the individual forecast models.

E. DISCUSSION

In viewing the case studies, the blending of individual model forecasts in the regressed and equally weighted ensemble forecasts is clearly evident. Additionally, the differences between the regressed and equally weighted forecasts give some

indication of the unequal weighting being applied to the individual forecasts in the blending by the regression method. As expected, anomalous forecast differences were identified which were not identified by the RMS error statistics compiled for the complete data set.

Overall, the regressed forecasts provided the greatest reduction in forecast error when compared with each of the other forecasts. Although a few cases did occur where individual model forecasts verified more accurately, the regressed forecasts displayed a consistently improved performance over the study domain. Given the model blending present in the ensemble forecasts, the regressed forecast exhibited more stability in terms of consistently lower error magnitudes than the individual model forecasts. Individually, the single model forecasts were more likely to display an occasional "bust" in a forecast. In the event that one model forecast performed significantly worse than the others, the performance of the regressed forecast was somewhat degraded due to the contribution from the weak forecast but would retain some stability by lowering the overall error through the blending of the entire set of individual model forecasts. In cases where all of the individual forecasts performed poorly, the regressed forecast also generally performed poorly. An exception to this case could occur if the weak performance of all contributing forecasts could be attributed

to systematic error which could be identified and reduced by the regressed forecast.

The characteristics of the errors in the regressed forecasts generally reflected the error characteristics of the ensemble contributors. If a majority of the individual model forecasts exhibited a strong phase vice amplitude error for a synoptic feature, then the regressed forecast often also reflected the phase error. However, the blended or statistical nature of the regressed forecast generally allowed the product to moderate or reduce the error. Furthermore, the ability of the regressed forecast to identify and reduce a systematic contribution to the error, allowed the regressed forecast to achieve gains over the equally weighted ensemble blend.

The regressed forecasts seemed to provide more consistent gains over the other forecast products at the surface rather than at 500 hPa. This may be attributed to the ability of the regressed method to identify a greater degree of systematic error in features at the surface for this data set. The limited number of case studies investigated, however, may not justify a trend of this nature.

The regressed forecasts did particularly well at the surface in forecasting lee cyclone features. This is perhaps attributable to the fact that many models have systematic errors when forecasting these features which the regression method is able to identify and reduce.

F. FORECAST DIVERGENCE AS A PREDICTOR OF FORECAST SKILL

As shown in the background section of this study, the ability to predict forecast skill has received much attention in recent years. The divergence or spread of individual members of an ensemble forecast has been shown to have significant correlation with forecast error. Although not the primary focus of this study, the divergence of individual forecasts was calculated and plotted for the case study events.

As an example, a plot of the forecast divergence for the 500-hPa forecasts on 7 September is shown in Fig. 37. In order to use forecast divergence as a predictor of forecast skill, a high degree of correlation is desired between areas of maximum spread in the forecasts and high error regions. A high degree of correlation in these values did exist between the example divergence plot in Fig. 37 and the corresponding error plot of the regressed forecast in Fig. 26b. Although the high correlation in these plots was not as consistent in other cases that were investigated, the potential of using forecast divergence as a forecast skill predictor in this method is clearly present and requires further study.

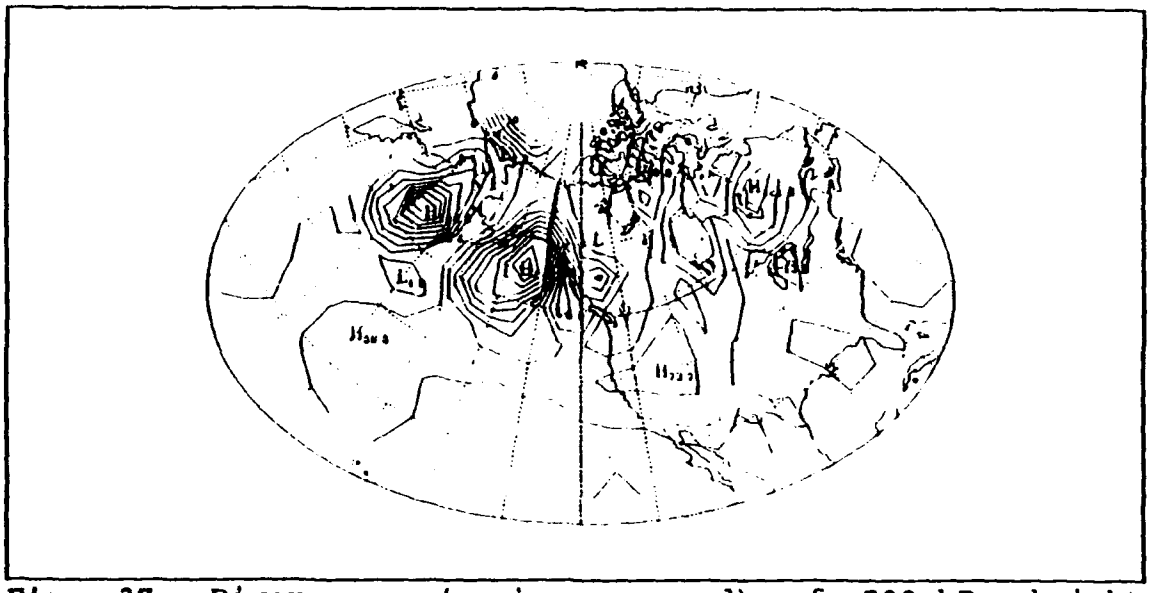

Fig. 37. Divergence (maximum spread) of 500-hPa height individual model forecasts at valid time 1200 UTC 7 September 1992 (contour interval 20 m).

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

The regression method of ensemble forecasting has been applied with success in this study. Both RMS error and case study verifications have been used to show that the mean forecast error can be reduced through the statistical blending of individual model forecasts.

The regressed and equally weighted forecast models both showed improvement over the individual models in the 72-h forecast period. These ensemble forecasts showed even greater improvement over selected individual model forecasts in the extended-range forecast period. The regressed and equally weighted forecasts performed comparably in the RMS error statistics, however, the case studies indicated that additional gain from systematic error reduction is achieved by the regression method.

When applied to the dependent data set, the regressed forecasts clearly outperformed the equally weighted forecasts as expected. The lack of an independent data set from the summer season necessitated the use of fall and winter season data as the independent data set. For this independent data set, the relative error reduction was less than for the dependent data set. The results for the dependent data set

highlight the possibility that further gains might be achieved when the regression method is applied to a corresponding summer season of independent data.

B. RECOMMENDATIONS

Although favorable results were achieved from the single-year data set used in this study, the derivation of regression equations from a multi-year data set would likely improve the statistical performance of the regression method. Additionally, regression statistics developed from multiple years of seasonal data should show further gains by identifying seasonal characteristics of systematic error in the forecast models. Seasonal regression equations could be developed for application during, for example, the summer or winter periods. Similar seasonal statistical applications such as the Model Output Statistics have been applied in this manner with success. By increasing the size of the data set over time, however, there is an increased chance that modelling centers may incorporate model changes or enhancements in the forecast models which can affect the regression statistics.

Further study could be completed to develop the regression method for application to particular geographical regions, synoptic features, or even individual wavenumbers where characteristic forecast biases may be identified in forecast model performance. Increased focus of the method in this

fashion would allow the regression method to exercise its potential in identifying and reducing systematic error.

While only three forecast models were employed in the ensemble forecast blend in this study, a greater number of models could be incorporated in the regressed model development. Increasing the number of input models could provide even greater consistency to the method and possibly improve its overall performance.

Application of the regression method to even longer forecast periods such as ten days would be useful given the increased variation that occurs in individual forecasts over extended forecast periods. As shown in this study, ensemble forecast performance over selected individual products was significantly improved over a five-day forecast period.

Although use of the ECMWF analyses as verification products in this study did not appear to significantly degrade the verification of the NMC or FNOC forecasts, further study using another model's analyses in verification would be useful to establish the robustness of this technique. Another option would be to use an averaged analysis from the three model products in the verification studies.

LIST OF REFERENCES

Bennett, A. F., and L. M. Leslie, 1981: Statistical correction of the Australian region primitive equation model. *Mon. Wea. Rev.*, **109**, 453-462.

Chen, W. Y., 1989: Another approach to forecasting forecast skill. *Mon. Wea. Rev.*, **117**, 427-435.

Dalcher, A., E. Kalnay and R. N. Hoffman, 1988: Medium range lagged average forecasts. *Mon. Wea. Rev.*, **116**, 402-416.

Dalcher, A., E. Kalnay, R. Livezey and R. N. Hoffman, 1985: Medium range lagged average forecasts. *Preprints from the Ninth AMS Conference on Probability and Statistics in Atmospheric Sciences*, 130-136.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203-1211.

Glowacki, T., 1988: Statistical corrections to dynamical model predictions. *Mon. Wea. Rev.*, **116**, 2614-2627.

Hoffman, R. N., and E. Kalnay, 1983: Lagged-average forecasting. *Tellus*, **35A**, 100-118.

Kalnay, E., and A. Dalcher, 1987: Forecasting forecast skill. *Mon. Wea. Rev.*, **115**, 349-356.

Leith, C., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409-418.

Leslie, L. M., K. Fraedrich and T. J. Glowacki, 1989: Forecasting the skill of a regional numerical weather prediction model. *Mon. Wea. Rev.*, 117, 550-557.

Leslie, L. M., and G. J. Holland, 1991: Predicting regional forecast skill using single and ensemble forecast techniques. *Mon. Wea. Rev.*, **119**, 425-436.

Murphy, J. M., 1990: Assessment of the practical utility of extended range ensemble forecasts. *Quart. J. R. Met. Soc.*, **116**, 89-125.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria VA 22304-6145	2
2.	Library, Code 052 Naval Postgraduate School Monterey CA 93943-5002	2
3.	Chairman (Code OC/Co) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-50003.	1
4.	Chairman (Code MR/Hy) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-5000	1
5.	Professor Wendell A. Nuss (Code MR/Nu) Department of Meteorology Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Professor Carlyle H. Wash (Code MR/Wx) Department of Oceanography Naval Postgraduate School Monterey, CA 93943-5000	1
7.	Commander Naval Oceanography Command Stennis Space Center MS 39529-5000	1
8.	Commanding Officer Naval Oceanographic Office Stennis Space Center MS 39529-5001	1
9.	Commanding Officer Fleet Numerical Oceanography Center Monterey, CA 93943-5005	1

10. Commanding Officer 1
Naval Oceanographic and Atmospheric
Research Laboratory
Stennis Space Center
MS 39529-5004

11. Superintendent 1
Naval Research Laboratory
NPS Annex, Code 7500
Monterey, CA 93943-5006

12. Chief of Naval Research 1
800 N. Quincy Street
Arlington, VA 22217

13. LCDR David W. Titley 1
Commander, Carrier Group SIX
Unit 60103
FPO AA 34099-4306